The Astrophysical Journal, 685:705–724, 2008 October 1

A *HUBBLE* AND *SPITZER* SPACE TELESCOPE SURVEY FOR GRAVITATIONALLY LENSED GALAXIES: FURTHER EVIDENCE FOR A SIGNIFICANT POPULATION OF LOW-LUMINOSITY GALAXIES BEYOND $z = 7$

Johan Richard,[1] Daniel P. Stark,[1] Richard S. Ellis,[1] Matthew R. George,[1,2] Eiichi Egami,[3] Jean-Paul Kneib,[1,4] and Graham P. Smith[1,5]

Received 2008 March 31; accepted 2008 June 24

ABSTRACT

We present the results of a systematic search for gravitationally lensed continuum Lyman break "dropouts" beyond a redshift 7 conducted via very deep imaging through six foreground clusters undertaken with the *Hubble* and *Spitzer* space telescopes. The survey has yielded 10 z-band and two J-band dropout candidates to photometric limits of $J_{110} \simeq 26.2$ AB (5 σ). Taking into account the magnifications afforded by our clusters (1–4 mag), we probe the presence of $z > 7$ sources to unlensed limits of $J_{110} \simeq 30$ AB, fainter than those charted in the Hubble Ultradeep Field. To verify the fidelity of our candidates we conduct a number of tests for instrumental effects which would lead to spurious detections, and carefully evaluate the likelihood of foreground contamination by considering photometric uncertainties in the dropout signature, the upper limits from stacked IRAC data and the statistics of multiply imaged sources. Overall, we conclude that we can expect about half of our sample of z-band dropouts to be at high redshift. An ambitious infrared spectroscopic campaign undertaken with the NIRSPEC spectrograph at the W. M. Keck Observatory for seven of the most promising candidates failed to detect any Lyα emission highlighting the challenge of making further progress in this field. While the volume density of high-redshift sources will likely remain uncertain until more powerful facilities are available, our data provides the first potentially interesting constraints on the UV luminosity function at $z \simeq 7.5$ at intrinsically faint limits. We discuss the implications of our results in the context of the hypothesis that the bulk of the reionizing photons in the era $7 < z < 12$ arise in low-luminosity galaxies undetected by conventional surveys.

Subject headings: cosmology: observations — galaxies: evolution — galaxies: formation — galaxies: high-redshift — gravitational lensing

Online material: color figures

1. INTRODUCTION

Very little is currently known about the abundance and luminosity distribution of star-forming sources beyond $z \gtrsim 7$. The two principal techniques used to locate distant star-forming sources at lower redshifts, the Lyman break "dropout" technique (Bouwens et al. 2006) and the location of Lyα emitters (Kashikawa et al. 2006; Shimasaku et al. 2006), become challenged by the lower performance of infrared instruments. In addition, the likely sources are much fainter, particularly if an increasing fraction are subluminous as might be expected given mass assembly is at an early stage (Loeb & Barkana 2001; Choudhury & Ferrara 2007). Despite these hurdles, it seems reasonable to expect that there is an abundance of star-forming galaxies at these epochs. The improved measurement of the optical depth to electron scattering derived from temperature-polarization correlations in the microwave background (Komatsu et al. 2008) suggests reionization occurred around $z_{\rm reion} = 10.8 \pm 1.4$ assuming it happened instantaneously; more probably it proceeded over an extended redshift window $7 < z < 12$ (Spergel et al. 2007). Moreover, the detection of galaxies at $z \sim 6$ with significant stellar masses and mature stellar populations (Stark et al. 2007a; Eyles et al. 2007) and the ubiquity of ionized carbon in the intergalactic medium probed by the highest redshift QSOs (Songaila 2004; Ryan-Weber et al. 2006) together demand a significant amount of star formation at earlier times, possibly enough to cause reionization. Although uncertainties remain, these independent arguments strongly motivate the search for $z > 7$ star-forming sources.

Most of the early progress in this quest has been made through the publicly available deep *Hubble Space Telescope* (*HST*) images. The Great Observatories Origins Deep Survey (GOODS; Dickinson et al. 2003), and the Hubble Ultra Deep Field (Beckwith et al. 2006 UDF) and its associated parallel fields have been used to search for "dropouts" in the i_{775} (Bouwens et al. 2006; Bunker et al. 2004), z_{850} (Bouwens & Illingworth 2006; Bouwens et al. 2008) and J_{110} (Bouwens et al. 2005, 2008) bands, corresponding to effective source redshifts of $z \simeq 6$, 7.5, and 10. These studies found a highly uncertain number density of candidates, none of which has been confirmed spectroscopically at $z > 7$. However, taken at face value, the overall conclusion from these ultradeep images is that the declining abundance of luminous star-forming sources beyond $z \simeq 7$ is insufficient to account for reionization. Although there is no guarantee that star-forming sources did reionize the universe at $z \simeq 10$, a possible solution is that the bulk of the early star formation resides in an undetected population of intrinsically subluminous sources (Stark et al. 2007a).

Prior to the availability of the next generation of telescopes, gravitational lensing is an effective means to evaluate this hypothesis. Depending on the method, foreground massive clusters can provide a magnification boost of $\times$5–30 in flux (for

[1] Department of Astrophysics, California Institute of Technology, MS 105-24, Pasadena, CA 91125; johan@astro.caltech.edu.
[2] Current address: Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK.
[3] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721.
[4] Observatoire Astronomique de Marseille-Provence, Traverse du Siphon-BP 8, 13376 Marseille Cedex 12, France.
[5] School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

unresolved sources) or in size (for resolved sources). An analysis by Stark et al. (2007c) concluded that lensing surveys should be able to find ample candidates at $z \gtrsim 7$, permitting follow-up spectroscopy and detailed studies at sensitivity limits that would be unachievable otherwise. As pointed out by Broadhurst et al. (1995) this gain is offset by a reduction in the sky area surveyed, producing an overall increase or decrease in the number of lensed sources, depending on the slope of the luminosity function.

A first attempt at constraining the abundance of lensed dropout galaxies at $6 \lesssim z \lesssim 10$ was made by Richard et al. (2006) using deep ISAAC near-infrared images obtained at the ESO VLT. A number of faint [intrinsic $H(\mathrm{AB}) \sim 26$] candidates were identified in two clusters, where the *magnification factor* μ ranged from 0.4 to 2.5 mag. This analysis suggested a star formation rate density at $z \simeq 7.5 \sim \times 10$ higher than that derived by Bouwens et al. (2004). Deeper NICMOS images have failed to confirm some of these sources. The number densities of faint candidates are currently being revised using follow-up imaging and spectroscopy for a small fraction of the ISAAC field of view (J. Richard et al., in preparation). The deeper NICMOS imaging undertaken in the present study illustrates the difficulties in making progress beyond $z \simeq 7$ using ground-based facilities.

In a parallel effort, Stark et al. (2007b) concentrated on the much smaller regions of very high magnification ($\mu > 3$ mag) termed the "critical lines" of the lensing clusters. Using NIRSPEC on the Keck Telescope they undertook a "blind" spectroscopic search for lensed Lyα emitters in the redshift range $8.5 < z < 10.4$. Despite the very small volumes probed in this unique survey, six faint candidates emerged across nine clusters. Exhaustive follow-up imaging and spectroscopy has, so far, been unable to provide unambiguous confirmation of the nature of these sources. Taking into account the uncertainties, Stark et al. 2007b concluded that the abundance of low-luminosity emitters in this redshift window may exceed 0.2 Mpc^{-3}, suggestive of a major contribution of low-luminosity star-forming galaxies to cosmic reionization.

The caveats concerning this conclusion were discussed in detail by Stark et al. Although representing a unique search for early star-forming sources at limits well beyond those probed otherwise, the volumes addressed are modest and significantly affected by cosmic variance. The six candidates were found in only three of the nine lensing clusters; six clusters had no convincing candidates. Although Stark et al. were unable to prove, unambiguously, that the detected emission is due to Lyα, the null detection of associated lines was used as an argument for rejecting lower redshift emission for most of their candidates. Their conclusion that the bulk of the reionizing photons arose from low-luminosity ($\simeq 0.1\ M_{\odot}\ \mathrm{yr}^{-1}$) star-forming sources can be verified by this independent search for lensed *continuum* dropouts at $z > 7$.

In searching for faint lensed dropouts, the advantage of *HST* over a ground-based survey such as that undertaken by Richard et al. (2006) is considerable. The ACS and NICMOS cameras are much more sensitive, not only because of the reduced background level relative to that produced by the atmosphere and by ambient temperature optics, but also because the typical sources have angular sizes of $0.2''$ or less (Ellis et al. 2001). With similar exposure times, NICMOS can readily attain a depth of 26.5 AB mag, $\sim$1 mag deeper than the earlier VLT/ISAAC project. Viewed through a typical $z \sim 0.2$ cluster, the NICMOS field ($0.8' \times 0.8'$) closely matches the area of moderate to high magnification factors ($\mu \sim$ 1–4 mag). The effective increase in sensitivity to faint sources provided by the lensing magnification along the line of sight to each NICMOS field probes limits fainter than those in the UDF (Bouwens et al. 2008), albeit over a considerably reduced area.

Early studies of lensed dropouts with *HST* have served to illustrate the potential. Kneib et al. (2004) located a source at $z \simeq 6.8$ behind the cluster Abell 2218. This source forms a triply imaged system with two bright elongated arcs, easily recognized as morphologically similar in ACS and NICMOS images. Follow-up observations with *Spitzer* (Egami et al. 2005) provided improved constraints on the photometric redshift, stellar mass ($M \sim 10^9\ M_{\odot}$), and past star formation history. Very recently, a similar $z \simeq 7.6$ lensed source was found by Bradley et al. (2008), but this was not multiply imaged. A semianalytic analysis, empirically calibrated using the luminosity functions of Lyα emitters and dropout galaxies at $z \sim 5$–6 (Stark et al. 2007c) predicts that such a NICMOS/ACS lensed imaging survey should typically detect 0.5–1 sources per cluster in the redshift range $7.0 < z < 8.5$.

The present program represents the logical next step: a concerted effort to verify the hypothesis advocated by Stark et al. as well as its associated predictions via deep imaging of six lensing clusters with *HST* (GO 10504: PI: Ellis) and *Spitzer* (GO-20439, PI: E. Egami). The primary goal is to determine the abundance of intrinsically subluminous z and J_{110} dropouts, and to derive constraints on the possible contribution of low-luminosity sources to cosmic reionization, independently of Lyα searches in blind or narrowband surveys.

The paper is organized as follows. In § 2 we present the new *HST*/*Spitzer* and associated ground-based imaging observations and their data reduction. A discussion of the means of selecting the dropout candidates is presented alongside a catalog in § 3. Issues of completeness and contamination by spurious and foreground sources are discussed. We conclude that a significant fraction of our candidates are possibly at high redshift. We then describe a Keck spectroscopic follow-up campaign in § 4, which attempts to detect Lyα emission from some of the most promising candidates. In § 5 we discuss the UV luminosity function at $z \simeq 7.5$ and review the implications in the context of the possible contribution of low luminosity to cosmic reionization. Our conclusions are summarized in § 6.

All magnitudes given in this paper are standardized within the AB magnitude system (Oke 1974). We assume a flat universe and $(\Omega_m, \Omega_\Lambda, h) = (0.3, 0.7, 0.7)$ whenever necessary.

2. SURVEY STRATEGY, OBSERVATIONS, AND DATA REDUCTION

We begin by discussing the selection of lensing clusters, the imaging data sets we have secured to select the various dropout candidates and validate their high-redshift nature, and the image processing steps taken to produce photometric catalogs. In addition to the *HST* and *Spitzer* data sets which form the fulcrum of this study, we have added ground-based imaging in both the optical and K band. In general terms, the *HST* data acts as the primary source of dropout candidates and the associated data is used to constrain the likely redshift distribution.

2.1. *Lensing Cluster Sample*

Our primary criterion in selecting foreground clusters for this lensing survey is the value of the magnification factor expected for sources beyond $z \simeq 7$ and the uncertainty implied in this estimate based on an understanding of the mass model. We considered a number of massive clusters at $z \simeq 0.1$–0.5 with well-understood mass models capable of producing regions of strong magnification which match the imaging area of the *HST* cameras. Even though the area enclosed by the critical line for a $z \simeq 7$ source is sufficient for the design of the program, a precise mass model

TABLE 1
Lensing Cluster Sample

Cluster	R.A. (deg)	Decl. (deg)	z	Mass Model	$N_{\rm mult}(N_z)$[a]
Abell 2218	248.95625	66.21444	0.176	Elíasdóttir et al. (2007)	37 (26)
Abell 2219	250.08541	46.70833	0.226	Smith et al. (2005)	14 (6)
Abell 2390	328.40086	17.69603	0.228	Swinbank et al. (2006)	11 (5)
Abell 2667	357.91387	−26.08541	0.233	Covone et al. (2006)	10 (5)
Cl 0024+16	6.65122	17.16060	0.390	Kneib et al. (2003)	9 (5)
Cl 1358+62	209.96069	62.51808	0.328	Franx et al. (1997)	5 (3)

[a] $N_{\rm mult}$ is number of multiple images used in the lens model (N_z number with spectroscopic redshifts)

is a clear advantage in determining the magnification of a particular source, as well as for predicting accurately the location of any multiple images.

Our strategy parallels that discussed in some detail by Stark et al. Indeed, that Lyα critical line survey placed greater demands on the reliability of the cluster mass models as the magnifications are more extreme and positional uncertainties in the image plane are critical. Notwithstanding this challenge, Stark et al. found that the typical magnification uncertainties for their candidates were only $\simeq$30% and that errors in the critical line location were usually only $\pm 1''$ or so.

In the present survey, six clusters was considered the minimum number necessary to mitigate the effects of cosmic variance (Stark et al. 2007c) while being consistent with the observing time available. In considering the final tally of clusters, we included clusters which we have modeled with the latest version of the Lenstool[6] software (Jullo et al. 2007). This improved code provides a new Bayesian optimization method to derive error estimates for each optimized parameter. Following the discussion in Richard et al. (2007) and Limousin et al. (2007), this optimization can be used to compute error estimates on the individual magnification factors for each of our sources.

Wherever possible, we included clusters with usefully deep archival *HST* and *Spitzer* IRAC data. Deep optical (AB > 27.0) ACS and/or WFPC2 images, previously used for the identification of multiple images during the development of the mass models, allow low-redshift contaminants to be identified reliably. IRAC images at 3.6 and 4.5 μm, to the same 1 μJy sensitivity as our previous work in Abell 2218 Egami et al. (2005) are available as part of a *Spitzer* Lensing Cluster Survey (GTO 83, PI: G. Rieke) or publicly from the archive (GTO 64, PI: G. Fazio). Finally, we required that each of the selected clusters be visible from the northern hemisphere, in order to facilitate spectroscopic follow-up and further *K*-band imaging with the Keck and Subaru telescopes.

The six clusters satisfying the above criteria are presented in Table 1. Although the references cite the most recently published mass models, as described above, in each case we have utilized the available multiple images and their redshifts in improving these mass models using Lenstool. Four out of the six clusters are in common with the sample adopted by Stark et al.

2.2. *Hubble Space Telescope Data and Reduction*

Our large program with *HST* (GO 10504, PI: Ellis) comprised deep observations in the z_{850}, J_{100}, and H_{160} bands, using the Wide Field Camera of ACS and the NIC3 configuration of NICMOS. The region enclosed by the critical line for the putative $z > 7$ sources was typically covered by two NICMOS pointings per cluster, usually adjacent. This ensured a magnification gain of $\mu \sim 1$ to 4 mag, with a typical value $\mu \sim 2$ mag, throughout the NICMOS imaged field (Fig. 1). The total sky area covered by the NICMOS observations is 8.9 arcmin^2 for the six clusters.

The ACS images were reduced with the `multidrizzle` software (Koekemoer et al. 2002). This removes cosmic rays and bad pixels and combines the dithered frames to correct for camera distortion. The output pixel scale was fixed at $0.04''$and we used a `pixfrac` value of 0.8 for reducing the area of the input pixels. We made small corrections to the absolute astrometry to allow for ACS frames taken at different epochs (e.g., Abell 2218 in the F850LP band; see Table 2). These corrections were computed by correlating the catalog positions of bright objects detected in the overlapping regions.

Each single set of NICMOS observations consists of eight (in F110W) and 10 (F160W) frames of $\sim$1000 s, taken with the NIC3 camera using the SPARS64 or SPARS128 sampling sequences. A basic reduction was performed by adopting the procedures given in the NICMOS data reduction handbook.[7] Starting with the postcalibrated frames, bad pixels are flagged and rejected based on individual histograms, cosmic ray are rejected using the `LACOSMIC` (van Dokkum 2001) IRAF procedures, frame-to-frame shifts are measured using a cross correlation technique, and all

[6] Publicly available, see http://www.oamp.fr/cosmology/lenstool to download the latest version.

[7] See http://www.stsci.edu/hst/nicmos/documents/handbooks/handbooks/DataHandbookv7.

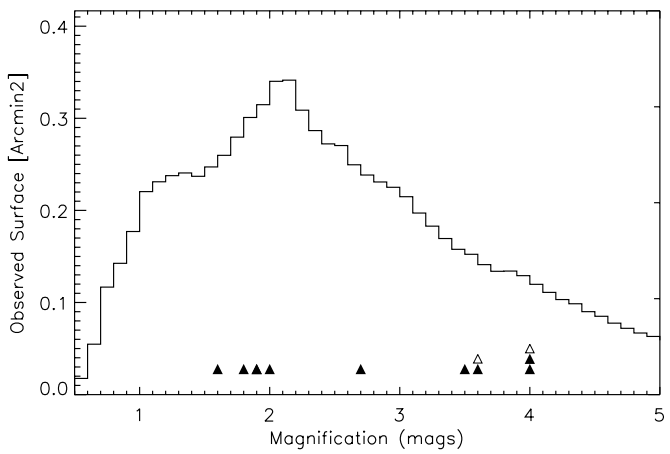


Fig. 1.—Distribution of magnification factors μ (in magnitudes) for the survey, as predicted from the mass models assuming a point source at very high redshift ($z > 7$). The vertical axis represents the observed surface area in each 0.1 mag (0.04 dex) magnification bin. The peak of this distribution indicates our typical magnification factor is $\sim$2.0 mag. Filled triangles mark the individual magnification factors for the 10 z dropout and open triangles that of the J dropout candidates.

TABLE 2
IMAGING DATA

Cluster	Filter	Program	Date	Exposure Time	Depth ($5\ \sigma$)	Completeness
A2218	ACS_{435}	9717	2004 Aug	7048	27.74	
	ACS_{475}	10325	2004 Aug	5640	27.95	
	ACS_{555}	9717	2004 Aug	7048	27.79	
	ACS_{625}	9717	2004 Aug	8386	27.93	
	ACS_{775}	10325	2004 Aug	9285	27.73	
	ACS_{850}	9292/9452/10325	2002 Apr, 2002 Aug, 2004 Aug	19630	27.32	
	$NIC3_{110}$	9452/10504	2003 Apr, 2005 Dec	8446	26.26	25.93
	$NIC3_{160}$	9452/10504	2003 Apr, 2005 Dec	10559	26.76	26.07
	$NIRC_K$		2002 Jul	7200	25.5	
	$NIRC_K$		2006 Jul	13620	25.9	
	$MOIRCS_K$		2007 May	18000	25.7	
	$IRAC_{3.6}$	83/20439	203 Dec, 2005 Oct, 2005 Dec	37700	25.5	
	$IRAC_{4.5}$	83/20439	2003 Dec, 2005 Oct, 2005 Dec	37700	25.5	
A2219	$WFPC2_{702}$	6488	1999 Aug	14400	27.00	
	ACS_{850}	10504	2006 Apr	8374	26.75	
	$NIC3_{110}$	10504	2006 May, 2006 Jun, 2007 Jun	9216	26.23	25.97
	$NIC3_{160}$	10504	2006 May, 2006 Jun, 2007 Jun	11519	26.73	26.14
	$NIRC_K$		2006 Jul	22980	26.3	
	$MOIRCS_K$		2006 Aug	17550	25.7	
	$IRAC_{3.6}$	83	2004 Feb, 2005 Mar	2400	23.9	
	$IRAC_{4.5}$	83	2004 Feb, 2005 Mar	2400	23.9	
A2390	$WFPC2_{555}$	5352	1994 Dec	8400	26.6	
	$WFPC2_{814}$	5352	1994 Dec	8400	26.2	
	ACS_{850}	9292/10504	2002 May, 2006 May	8847	26.82	
	$NIC3_{110}$	10504	2006 Jun, 2006, Jul, 2007 Jun	9470	26.27	25.89
	$NIC3_{160}$	10504	2006 Jun, 2006 Jul, 2007 Jun	11839	26.54	26.06
	$MOIRCS_K$		2007 May	15900	25.6	
	$IRAC_{3.6}$	83	2004 Jun, 2004 Nov	2400	23.9	
	$IRAC_{4.5}$	83	2004 Jun, 2004 Nov	2400	23.9	
A2667	$WFPC2_{450}$	8882	2001 Oct	9600	26.26	
	$WFPC2_{606}$	8882	2001 Oct	4000	26.94	
	$WFPC2_{814}$	8882	2001 Oct	4000	26.11	
	ACS_{850}	10504	2006 Jul	8765	26.70	
	$NIC3_{110}$	10504	2006 Aug	9343	26.22	25.93
	$NIC3_{160}$	10504	2006 Aug	11711	26.51	26.01
	$IRAC_{3.6}$	83	2003 Dec	2400	23.9	
	$IRAC_{4.5}$	83	2003 Dec	2400	23.9	
CL 0024	ACS_{435}	10325	2004 Nov	6435	27.67	
	ACS_{475}	10325	2004 Nov	5072	27.81	
	ACS_{555}	10325	2004 Nov	5072	27.47	
	ACS_{625}	10325	2004 Nov	8971	27.75	
	ACS_{775}	10325	2004 Nov	10144	27.67	
	ACS_{850}	10325	2004 Nov	16328	27.28	
	$NIC3_{110}$	10504	2006 Aug	9472	26.20	25.9
	$NIC3_{160}$	10504	2006 Aug	11840	26.60	26.0
	$MOIRCS_K$		2006 Aug	21600	25.7	
	$IRAC_{3.6}$	64	2003 Dec	2400	23.9	
	$IRAC_{4.5}$	64	2003 Dec	3600	24.1	
CL 1358	ACS_{435}	9717	2004 Apr, 2004 May	5440	27.70	
	ACS_{475}	9717	2004 Apr, 2004 May	5470	27.96	
	ACS_{625}	9717	2004 Apr, 2004 May	6800	27.77	
	ACS_{775}	9717	2004 Apr, 2004 May	10144	27.49	
	ACS_{850}	9717	2004 Apr, 2004 May	16328	27.13	
	$NIC3_{110}$	10504	2005 Dec	9216	26.34	25.92
	$NIC3_{160}$	10504	2005 Dec	11519	26.60	26.10
	$IRAC_{3.6}$	83	2004 Jan, 2005 Jun	2400	23.9	
	$IRAC_{4.5}$	83	2004 Jan, 2005 Jun	2400	23.9	

For a given cluster each entry presents the instrument, filter, *HST*/*Spitzer* program ID, date of observation, exposure time, and final image quality (depth and completeness).

frames are drizzled onto a NICMOS pixel scale (0.2″) to produce an initial reduced image. This then serves as a comparison for each individual frame so that deviant pixels can be flagged to improve image quality in a second drizzling operation, this time undertaken with an image pixel of 0.1″, to obtain a better sampling.

These initial images reveal a number of cosmetic effects (bias and flat residuals, bad columns and bad pixels, quadrant-to-quadrant variations, background variations) that led to a second stage of image reduction. We used an improved pixel mask to flag several bad columns and bad pixels close to the frame edges, and we examined each frame individually in order to remove bias and flat residuals and quadrant-to-quadrant variations. Finally, we subtracted a smoothed background obtained by averaging the frames of all the observed clusters, masking every pixel lying on a physical object. In order to combine all NICMOS observations of a given cluster, usually taken at different epochs and with slight variations in the sky orientation, the individually reduced images were registered onto the wider field ACS data prior to the final drizzling procedure and combination into a NICMOS mosaicked image.

The NICMOS data acts as the primary basis for selecting our dropout candidates in association with nondetections in very deep optical data. In addition to our own ACS data, we reduced deep ACS and WFPC2 images from the archive for each cluster in our sample (Table 2). As with the ACS F850LP data undertaken in our own program, we reduced these data using the IRAF procedures `multidrizzle` and `drizzle` as discussed above.

2.3. *Ground-based Data and Reduction*

As mentioned, optical nondetection to deep limits is a key necessity in considering the validity of our dropout candidates. Most clusters have complete optical coverage with *HST* (from ~400 to ~850 nm) useful for this purpose. However, Abell 2219 has F702W only, and Abell 2390 has F555W and F814W only. For these two clusters, additional deep ($R \sim 26.7$ and $I \sim 25.6$ at 5 σ in 1.2″diameter apertures) ground-based images taken with the CFH12k camera on CFHT (Bardeau et al. 2007) were also examined to check for nondetections at the locations of the dropouts.

Ground-based photometry in the K band can likewise provide additional information for z dropouts, improving the photometric redshift estimates, and for the reliability of J dropouts where otherwise only a single color would be available (see § 3.1). Although a challenging undertaking given the depth of the ACS and NICMOS data, we conducted various K-band imaging observations of clusters using the Keck and Subaru telescopes.

The Near Infrared Camera (NIRC) on Keck I was used on 2006 July to observe the central regions of the clusters Abell 2218 and Abell 2219. The square field of NIRC (38″a side) is slightly smaller than that of NICMOS; therefore' we concentrated these observations on our best z- and J-band dropouts in each cluster (see § 3.1). The seeing was stable and in the range 0.5″–0.6″. We used dithered exposures of six co-adds ×10 s exposure time each. The NIRC pointing in Abell 2218 was partially covered, in the region of one of the dropouts, by previous NIRC observations obtained in 2002 July 22 and 23 (A. Blain & N. Reddy 2003, private communication).

The Multi-Object InfraRed Camera and Spectrograph (MOIRCS; Ichikawa et al. 2006) at Subaru was also used during two observing campaigns, in 2006 August and 2007 May. The larger field of view ($7' \times 4'$) of this camera ensures a complete coverage of the NICMOS, ACS, and IRAC data in each cluster. The four clusters in common with the Stark et al. sample (Abell 2218, 2219, and 2390, and CL 0024) were imaged for ~5 hr each under very good seeing conditions (0.3″–0.4″) using dithered exposures of 50 s duration. We used the MCSRED software package[8] to perform the flat-fielding, sky subtraction, distortion correction, and mosaicking of individual images. The MCSRED package also includes MOIRCS-specific tasks which correct for quadrant shifts and to fit sky residuals. The depth reached by the MOIRCS imaging data ($K_{AB} \sim 26.1$) in Abell 2218 and Abell 2219 is similar or deeper than the NIRC observations of the same fields. Therefore, we used the NIRC images as independent check for the MOIRCS photometric measurements, performed under better seeing conditions.

2.4. *Spitzer Data and Reduction*

Each cluster has been observed at 3.6, 4.5, 5.8, and 8.0 μm using the Infrared Array Camera (IRAC; Fazio et al. 2004) on board the *Spitzer Space Telescope* (Werner et al. 2004). In this paper we only discuss the first two channels (3.6 and 4.5 μm), where the depth achieved is potentially useful in comparison with our ACS and NICMOS data. Each of the IRAC channels uses a separate detector array, and the 3.6 μm channel ($\lambda_c = 3.56$ μm; $\Delta\lambda = 0.75$ μm) and 4.5 μm channel ($\lambda_c = 4.52$ μm; $\Delta\lambda = 1.01$ μm) use 256×256 InSb arrays with a pixel scale of 1.″2 pixel^{-1}, producing a field of view of $5.2' \times 5.2'$. A frame time of 200 s was used with the small-step cycling dither pattern initially, but this was later changed to 100 s with the medium-step cycling pattern for a better removal of cosmic rays and other artifacts.

Most of the IRAC data come from the *Spitzer* GTO program PID: 83 (PI: G. Rieke). The total integration time is 2400 s channel^{-1}, usually obtained at two different epochs. CL 0024 was observed as part of another GTO program PID: 64 (PI: G. Fazio) with the integration times of 2400 s at 3.6 μm and 3600 s at 4.5 μm. Abell 2218 was the subject of a deeper GO campaign (PID: 20439; PI: Egami), and was observed for ~10 hr per channel, split into six separate observations (AORs). When these GO data are combined with those of the GTO program (PID: 83), the resultant total integration time is 37,700 s per channel. The corresponding depth (25.5 AB) in this particular cluster is much closer to the magnitude of our high-redshift candidates. (see Table 2 for a summary of the data).

In producing the IRAC images, we started with the basic calibrated data (BCD) of each individual frame produced by the SSC pipeline, and combined them using a custom IDL mosaicking routine as presented in Egami et al. (2005). The final pixel size of the IRAC mosaics is 0.6″ pixel^{-1}, half that of the original data. A conservative estimate for the absolute calibration uncertainty is 10%.

2.5. *Foreground Subtraction*

Although the magnification afforded by lensing clusters offers a unique gain in probing the distant universe, the central regions are dominated by bright, extended spheroidal galaxies which obstruct and whose light increases the background level. More importantly, it also affects the photometry and color measurements of any underlying fainter source. Fortunately, the majority of these galaxies are also good light deflectors, their morphology is usually regular and their light distribution can be accurately modeled by a sequence of elliptical isophotes. Since the major contributor to this deficiency is the brightest cluster galaxy (BCG), we have modeled and subtracted its light for each cluster in the ACS (I_{850} band) and NICMOS images using the IRAF task `ellipse`. Apart from a small region (~1.0″ radius) around the core of the removed BCG, this frees the image from the majority of the contamination

[8] Available from http://www.naoj.org/staff/ichi/MCSRED/mcsred.html.

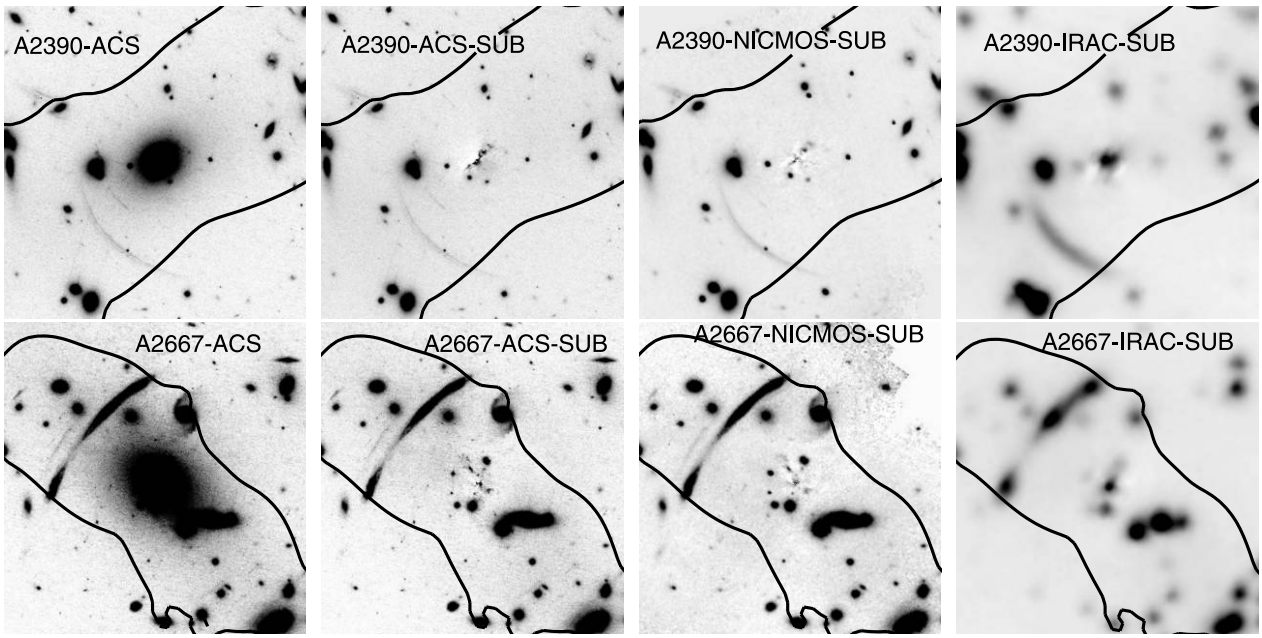


FIG. 2.—Example of BCG subtraction in the clusters A2390 and A2667. Each panel is 45″ on a side. *Left to right:* ACS image without subtraction, BCG-subtracted images in ACS, NICMOS, and IRAC, respectively. Subtraction residuals appear within ∼1″ of the center in the ACS and NICMOS case, ∼2″ for IRAC. The critical line for a high-redshift source is show by the black curve.

in highly magnified regions close to the critical line, and aids in the detection of underlying sources (see Fig. 2).

Foreground subtraction is more challenging for the IRAC images because of the much coarser PSF. The smaller number of available pixels per galaxy prevents the use of the `ellipse` modeling technique discussed above. Instead we adopted the `ellipse` model of the BCG derived from the NICMOS data and, on the assumption that the spectral energy distribution is identical across the galaxy at these long wavelengths, convolved this model with the IRAC PSF derived by stacking ∼50 bright isolated point sources. This model of the BCG was subtracted after an appropriate scaling factor. We found this procedure to be very effective, with residuals from the subtraction confined within a 2.0″ radius region (Fig. 2).

2.6. *Final Photometry*

The primary filters we use for identifying the near-infrared dropouts are the ACS z_{850} and NICMOS J_{110}, H_{160} bands available for each cluster. The bulk of the rest-frame UV flux is contained in the two reddest filters at high redshift, so we require all sources to be detected in both. The signal-to-noise ratio is improved by combining the J and H images, once normalized to a similar noise level, to form a single $\langle J + H \rangle$ detection image. This provides a more accurate measurement of the centroid and geometrical parameters of these objects,

We use the ''double-image'' mode of the SExtractor package Bertin & Arnouts (1996) to detect objects and compute magnitudes within a 0.6″ (NICMOS images) or 0.3″ (ACS images) diameter aperture. Corrections to total magnitudes, assuming a point source, were estimated using bright isolated unsaturated stars. The corresponding values are 0.3 and 0.6 mag for ACS and NICMOS, respectively.

The drizzling procedure used in the *HST* reductions, while conserving flux, does introduce correlations between neighboring pixels and hence unreliable error estimates (Casertano 2000). We applied their equation (A2) to the SExtractor photometric errors. Dithered exposures also introduce a varying effective exposure time across the mosaic, this effect is clearly apparent close to the boundaries of the field. We constructed a corresponding weight map, proportional to the effective integration time at each pixel position in the detection image, and used it as an input parameter of SExtractor. This weight map is normalized by the flat-field response of the detector to ensure a source detection at a uniform noise level across each image.

The background noise level σ was measured for the ACS and NICMOS images in order to estimate the achievable photometric limits in each band. The 5 σ values are reported in Table 2 as total magnitudes. We computed the completeness limits in each NICMOS filter used for source detection by adding artificial unresolved sources in the magnitude range 23-28 AB. Such sources were added 1000 times at 30 different random locations on the image and then extracted using the same photometric techniques as applied to the science image. In this manner we compute the point source completeness as a function of source magnitude. Only blank regions were chosen for this exercise, defined after application of a 5 σ threshold to mask pixels associated with bright objects. The 50% completeness values derived in this way are listed for each cluster and band in Table 2. Typically, our NICMOS data is 50% complete to $J_{110} \simeq 25.9$ AB and to $H_{160} \simeq 26.05$ AB. For each eventual dropout candidate (§ 3.3), we will assign a completeness factor, $S_{\rm comp}$, based on its magnitude, that will be taken into account in estimating the high-redshift luminosity function.

Our NICMOS survey reaches limits of typically 26.2 and 26.5 AB in J and H, respectively, over 8.9 arcmin^2. For comparison, the UDF limits are 27.8 AB over 7 arcmin^2 (Bouwens et al. 2004). However, the magnification provided by the foreground clusters (Fig. 1) enables us to reach unlensed source magnitudes (assuming $z \sim 7$) of 28-30 AB over a total area of 0.1–1 arcmin^2.

3. HIGH-REDSHIFT CANDIDATES

We now turn to the selection of our high-redshift candidates. The primary selection is based on the $z_{850} - J_{110}$ color for the

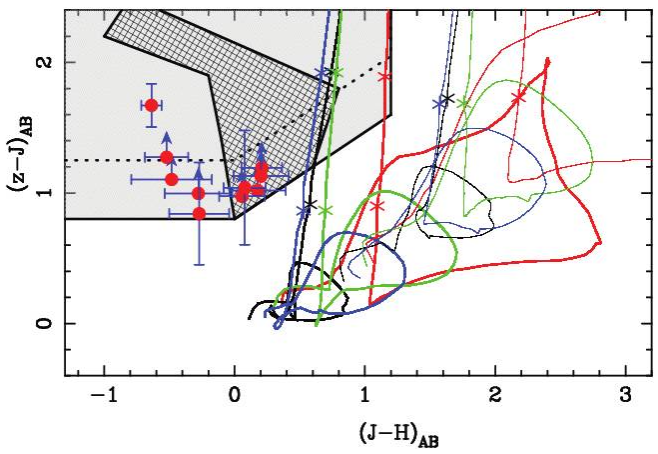


FIG. 3.—Color-color diagram used for selecting high-redshift z-band dropout candidates. The final set of candidates is presented in red (see § 3.3). Color tracks represent predicted colors of Hubble sequence galaxies (*colored tracks*; Coleman et al. 1980; Kinney et al. 1996). Thick tracks assume no extinction, thin lines show the effect of including a selective extinction of $A_V = 1.0$ mag. The observed location of L and T dwarfs is shown as a cross-hatched region (Knapp et al. 2004). Two possible $z - J$ color selections (>0.8 and >1.25) are shown (see text for details).

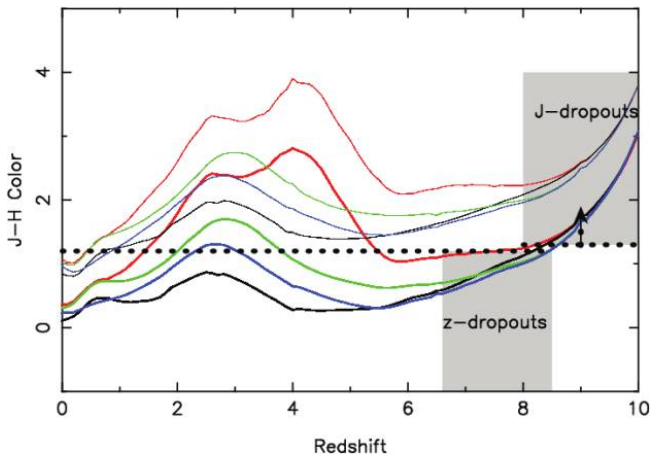


FIG. 4.—Optimizing the selection of high-redshift z-band and J-band dropout candidates using near-infrared colors (*dotted lines*) as compared to the expected colors of Hubble sequence galaxies (as in Fig. 3).

z dropouts at redshifts $z \simeq 7$–8, and the $J_{110} - H_{160}$ color for the J-band dropouts at redshifts $z \simeq 8$–10. This section has two components. First we discuss the optimum color criteria for dropout selection, the degree of completeness and issues of possible contamination from spurious artifacts at these faint limits. In this way we establish a robust set of candidates whose likelihood of being at high redshift we then assess in the second part of this section using additional criteria including their photometry at other wavelengths and location with respect to the cluster mass model. The catalog of candidates is summarized in § 3.3.

3.1. *Photometric Selection and Completeness*

The primary concern in selecting high-redshift dropouts from photometric data alone is the issue of contamination from lower redshift objects, including $z \simeq 2$ early-type galaxies, dust-reddened objects over a wide redshift range or low-mass Galactic stars with deep molecular absorption bands. Figure 3 illustrates the problem. A single color-cut fails to isolate z-band dropouts from a variety of $z \simeq 2$–4 sources and the confusion is worse for J-band dropouts (Fig. 4).

This problem has formed the basis of much discussion in the literature. For the z-band dropouts, contamination can be reduced by considering a second color (Stanway et al. 2004), of which the most useful with NICMOS data is $J_{110} - H_{160}$ (Bouwens et al. 2004). Star-forming galaxies at high z should display a prominent discontinuity in $z - J$ while remaining blue in $J - H$, as illustrated in Figure 3. The first issue we address is the optimum cut in both colors, on which depend both the redshift range explored and the amount of contamination by lower redshift objects.

Figure 3 shows color-redshift tracks for various galaxy classes (and also includes, for convenience, the colors of our eventual candidates discussed in § 3.3). The location of these tracks suggests the following prescription for selecting sources with redshifts $6.8 \lesssim z \lesssim 8.0$:

$$(z_{850} - J_{110}) > 0.8; \quad (z_{850} - J_{110}) > 0.66;$$
$$(J_{110} - H_{160}) + 0.8; \quad (J_{110} - H_{160}) < 1.2.$$

The baseline $(z - J) > 0.8$ color selection we adopt above is similar to that used by Bouwens et al. (2004, 2008). However, as the photometric errors for our candidates are typically 0.2–0.3 mag, the probability of low-redshift contamination remains significant in the range $0.8 < (z - J) < 1.1$. We also explore the use of a more restrictive color cut $(z - J) > 1.25$, close to the criterion used by Bouwens & Illingworth (2006) ($(z - J) > 1.3$), also shown on Figure 3.

Application of this color selection reveals 10 possible candidates of which only two satisfy the more rigorous $(z - J)_{AB} > 1.25$ color cut. We discuss the merits of each of these candidates in more detail in the following sections. A striking feature is that all have $J - H$ colors bluer than predicted by the redshifting of local spectral energy distributions. Similar claims for blue rest-frame colors have been made for i-band dropouts (Stanway et al. 2004).

Concerning the J dropouts, corresponding to the redshift range $8.0 \lesssim z \lesssim 10.5$, cuts of $(J - H) > 1.8$ and $(J - H) > 1.3$ were adopted, respectively, by Bouwens et al. (2005) and Bouwens et al. (2008). In their shallower, larger area survey aimed at locating luminous J dropout candidates, Henry et al. (2007) adopted a more restrictive $(J - H) > 2.5$ cut. The large variation in these color cuts reflects the differing depths of optical exclusion in the various samples. The sample selected by Henry et al. is limited to NICMOS-detected sources in two bands only, with no deep supporting optical observations. Therefore, despite the apparent stringent color cut in $(J - H)$, it is still more likely to suffer from contamination by lower redshift sources. Noting our deep optical data, we adopted a $(J - H) > 1.3$ cut, revealing two candidates (Fig. 4). Neither would satisfy the $(J - H) > 1.8$ cut.

We now turn to the important question of how complete our likely dropout sample will be given our adopted magnitude limits with ACS and NICMOS. We can easily imagine that genuine dropouts will be missed because photometric errors will scatter points outside our selection region; likewise, lower redshift sources will be scattered into our color selection box.

Using the procedures adopted to determine the 50% magnitude completeness limits in § 2.6, we can estimate the color selection completeness and color selection contamination by introducing artificial objects with a range of J magnitudes (AB = 24–27) whose $z - J$ breaks have a range (0.6–1.2) and determining what fraction of objects which would lie outside of the box with perfect photometry but are scattered in, and what fraction of objects lying truly inside are scattered out. We assume a flat f_ν spectrum for the UV continuum between the J and H bands so

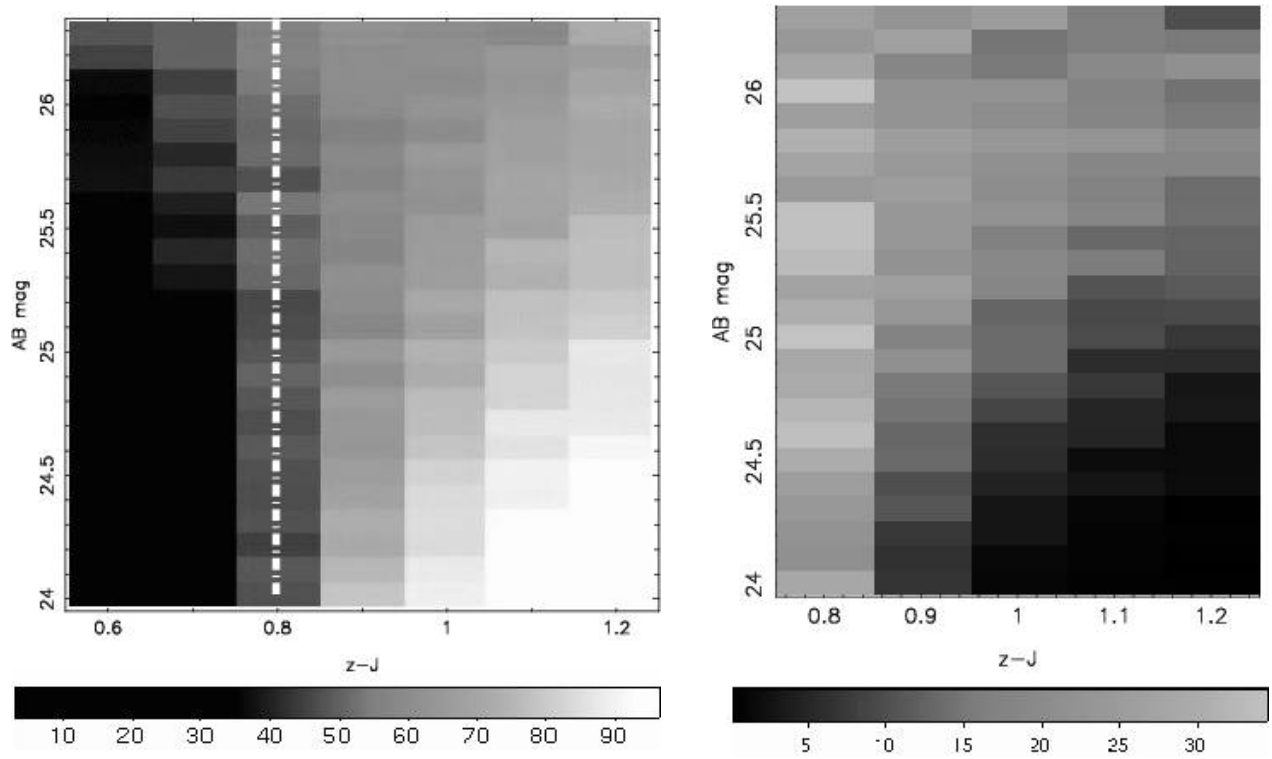


FIG. 5.—Testing the color-selection of z-band dropouts. *Left*: Completeness in the color-selection ($f_{\rm comp}$), as a function of the J or H magnitude (assuming a flat spectrum in AB) and the $(z - J)$ color break. Values represent the fraction (%) of simulated objects whose photometry satisfies our color-selection criteria as indicated in the color bar beneath. *Right*: As for the left panel but referring to the contamination fraction $f_{\rm cont}$. [*See the electronic edition of the Journal for a color version of this figure.*]

that, with reference to Figure 3, the problem becomes effectively one-dimensional.

The two panels in Figure 5 show the results of this test. The *selection completeness fraction* $f_{\rm comp}$ (*left*) represents the fraction of objects of a given magnitude and break color that we are able to recover in our selection. The *contamination fraction* $f_{\rm cont}$ (*right*) accounts for the fraction of objects with a lower break ($z - J < 0.8$) that are photometrically scattered in the color-color diagram so that their *observed* magnitude and color would allow them to enter our selection window. As with our magnitude completeness function, $S_{\rm comp}$ (§ 2.6), both of these correction factors will be used, for each candidate, to correctly estimate the true number density of objects having colors corresponding to $z > 6.8$ galaxies (i.e., a break $z - J$ above 0.8). Individual correction factors are reported in Table 3.

Examining Figure 5, as expected we find negligible difficulties for the brightest sources, but contamination and incompleteness become more troublesome at fainter magnitudes, depending on the $(z - J)$ color. We find that the selection completeness ranges from 50% to 95% and the contamination fraction is typically 15%.

3.2. *Verifying the Candidates*

The selection techniques discussed above yields a list of candidates for more careful examination. Here we discuss further

TABLE 3
HIGH-REDSHIFT CANDIDATE PHOTOMETRY

Candidate	R.A. (deg)	Decl. (deg)	z_{850} (AB)	J_{110W} (AB)	H_{160W} (AB)	K (AB)	μ (mag)	$S_{\rm comp}$	$f_{\rm cont}$	$f_{\rm comp}$
A2218-z1	248.9713	+66.2071	>27.32	26.1 ± 0.13	25.9 ± 0.11	>25.7	1.9	0.59	0.14	0.66
A2218-z2	248.9507	+66.2150	>27.32	26.2 ± 0.18	26.0 ± 0.11	>25.7	2.7	0.50	0.19	0.72
A2219-z1	250.0803	+46.7071	26.3 ± 0.15	24.7 ± 0.05	25.3 ± 0.06	>25.7	3.6	0.84	0.0	0.95
A2390-z1	328.4130	+17.6905	>26.82	25.5 ± 0.12	26.1 ± 0.12	>25.6	3.5	0.68	0.17	0.79
A2390-z2	328.4001	+17.6962	>26.82	25.8 ± 0.15	25.8 ± 0.10	>25.6	1.8	0.65	0.27	0.51
A2667-z1	357.9119	−26.0949	26.7 ± 0.36	25.9 ± 0.15	26.1 ± 0.18	N/A	1.6	0.59	0.34	0.51
A2667-z2	357.9153	−26.0826	26.7 ± 0.42	25.7 ± 0.12	25.6 ± 0.11	N/A	2.0	0.75	0.20	0.68
CL 1358-z1	209.9714	+62.5128	>27.33	26.3 ± 0.17	26.1 ± 0.12	N/A	1.9	0.43	0.15	0.63
CL 1358-z2	209.9521	+62.5108	>27.33	26.2 ± 0.13	26.7 ± 0.28	N/A	4.0	0.43	0.15	0.72
CL 1358-z3	209.9549	+62.5187	>27.33	26.3 ± 0.17	26.6 ± 0.19	N/A	4.0	0.35	0.14	0.66
A2219−j1	250.0900	+46.7040	>26.7	>26.3	25.0 ± 0.05	>25.7	4.0			
A2667-j1	357.9136	−26.0869	>26.7	>26.5	25.1 ± 0.08	N/A	3.6			

NOTES.—Total magnitudes and corresponding magnification assuming $z = 7.5$ for z-band dropouts and $z = 9.0$ for J-band dropouts. Each z dropout entry is followed by its observed completeness, selection contamination factor and selection completeness. Upper limits correspond to 5 σ.

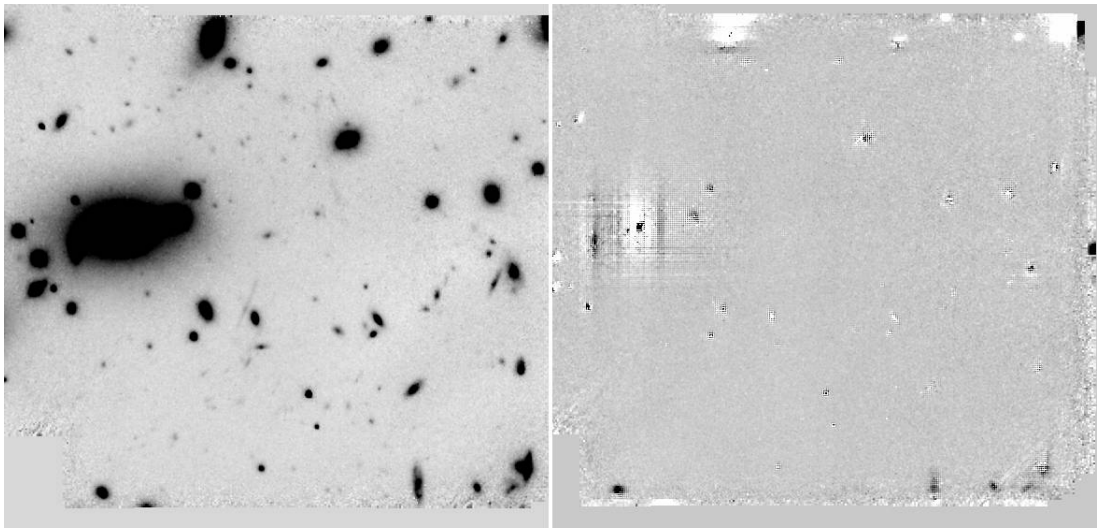


FIG. 6.—NICMOS image of the cluster CL 1358 (*left*) compared with a noise image (*right*) in order to estimate the fraction of spurious sources in our sample (see text for details). Except in the vicinity of the edges and central cores of bright objects (which are masked out by applying a simple threshold), the noise properties of the two images are very similar.

tests to determine the possibility that some might be spurious prior to establishing a catalog of genuine sources whose redshift distribution we explore using our additional photometric data.

3.2.1. *Spurious Detections*

As the signature of both our *z*- and *J*-band dropouts consists of a nondetection in the optical band, we must seriously consider the possibility of spurious detections in the NICMOS data. This is particularly the case for the *J* dropouts where only a single band is involved. An optical nondetection is defined as a implied flux lower than the photometric limit (5 σ in a 0.3″ diameter aperture, for ACS and WFPC2 data; see Table 2). These measurements were made using SExtractor in its "double-image" mode after the data was resampled and aligned to the NICMOS images. Further measures were made using the original multidrizzled images (see § 3.5).

A visual inspection was performed for each candidate in order to reject obviously spurious detections in the NICMOS images or false nondetections in the optical bands. The astrometric position of the candidates was used to perform this examination on the original images, to prevent biases arising from resampling and geometrical transformations. During this process, we rejected a number of candidates due to their proximity to the center of the removed BCG, or due to obvious contamination from very bright galaxy halos, both leading to noisier or biased photometric measurements.

Our photometric detection is based on the combination of 10 individual NICMOS images per pointing (four and six exposures in the J_{110} and H_{160} band, respectively). Because of a significant number of remaining hot and dead pixels in these individual frames, we investigated the fraction of spurious sources that would contaminate our photometric catalogs. To quantify this problem, we constructed a *noise image* for each cluster and near-infrared band whose purpose is to remove signals from all genuine sources while maintaining the same noise properties as the real data. This was done by subtracting in pairs the individual frames prior to eventual co-addition. We then applied our usual photometric detection software, using the same parameters as in the original images.

This noise image reveals residuals near the frame edges (due to the dithering process undertaken during the observations) and in the centers of the brightest objects, which were masked out in the detection process (Fig. 6). The affected area accounts, in total, for ~20% of the NICMOS field. By comparing the number of spurious sources detected in the noise image with the number of objects present in the original catalogs, we estimate the spurious fraction in the magnitude range of our candidates ($24.7 < J < 26.3$ and $25.3 < H < 26.7$) to vary from 4% to 18% from cluster to cluster, with an average value of 10%. Typically, therefore, we can expect around 90% of our candidates to be robust astronomical detections.

3.2.2. *Detector Remnance*

One specific worry, not addressed in the tests mentioned above, concerns the possibility of image persistence or ghosts appearing in the NICMOS frame during, or shortly after, an overexposure by a very bright source (Daou & Skinner 1997). The source producing the largest number of counts (~25,000 ADU) in our data set is a $J_{AB} = 17.9$ star in Abell 2390. We do not detect any measurable electronic ghost for this particular source, but persistence is seen at the level of a $J_{AB} = 24.9$ spurious source in an exposure taken 20 minutes later. Once averaged over the entire sequence (six exposures), this persistence corresponds to a $J_{AB} = 26.8$ source, i.e., fainter than our detection limit. We also verified, for each pair of successive exposures, that none of our candidates is coincidently located at the relevant position of similarly bright ($J < 19$) stars in all clusters.

Although NICMOS exposures from independent orbits are usually separated by a >30 minute delay, persistence might arise as a result of much brighter sources observed immediately prior to execution of our program. To eliminate this possibility, we located all preceding exposures and found no measurable persistence in our data, even in the case of a calibration program (GO 10726, PI: de Jong) aiming at measuring nonlinearity effects by repeatedly saturating the NIC3 detector.

3.3. *Catalog of Dropout Candidates*

As a result of our visual inspection of the initial candidate list selected according to the precepts of § 3.1, we emerge with 10 candidate *z* dropouts and two *J* dropouts. As seen in Figure 3, 2 of the 10 *z*-dropouts satisfy the most restrictive color-selection $(z - J) > 1.25$. The photometric measurements for these are

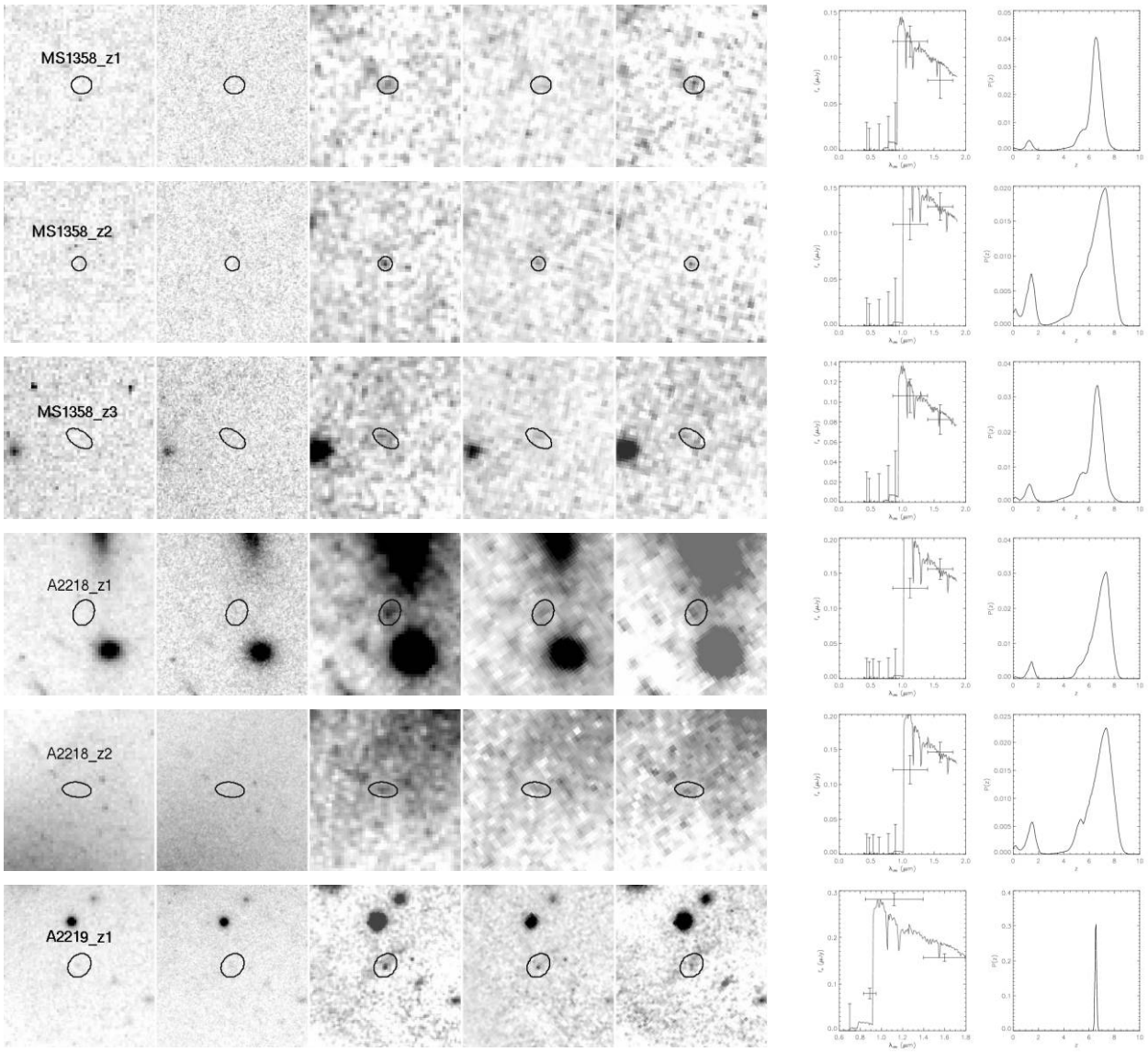


FIG. 7.—Snapshot images our final catalog of candidate dropouts. *Left to right*: *HST* images in the optical (ACS or WFPC2), in the ACS/F850LP (z) band, in the detection image (sum of F110W and F160W bands), in the NIC3/F110W and the NIC3/F160W images. To the right are the observed SED with the overplotted best-fit template for HyperZ, and the redshift probability distributions (see § 3.4.1. [*See the electronic edition of the Journal for a color version of this figure*.]

summarized in Table 3 and the relevant detection images are presented in Figure 7. Following the tests described above we can expect over 90% of these to be genuine astronomical sources.

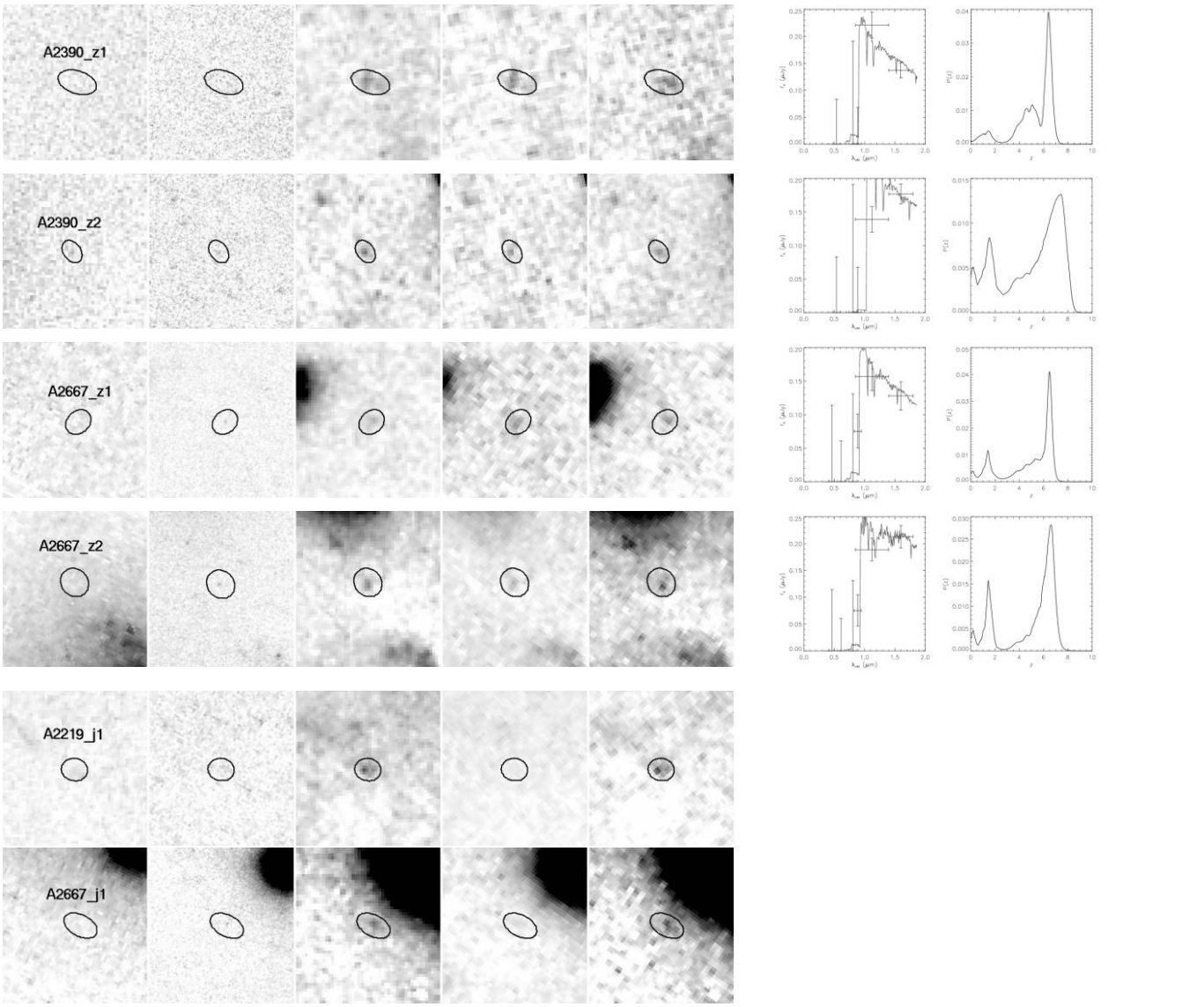
3.4. *Redshift Estimation*

We now turn to the important problem of contamination by lower redshift sources. We first use the spectral energy distribution (SED) of each candidate to estimate the individual photometric redshift. We then consider statistical arguments that can be applied to our entire candidate population.

3.4.1. *Photometric Redshifts*

As we have seen, the two color selection presented in § 3.1 enables us to select high-redshift galaxies with some confidence. However, we can use the multiwavelength data available for each source, including all upper limits arising from nondetections, in order to derive a *photometric redshift probability distribution*. To accomplish this, we used an updated version of the photometric redshift software HyperZ (Bolzonella et al. 2000). Best-fit redshift distributions were computed using a standard SED procedure with a variety of templates, including empirical data (Kinney et al. 1996; Coleman et al. 1980) and evolutionary synthesis models (Bruzual & Charlot 2003). We searched the redshift range $0 < z < 10$, while reddening was kept as a free parameter ranging between $A_V = 0$ and 2 mag, assuming the Calzetti et al. (2000) law. The effect of Lyman forest blanketing is included following the prescriptions of Madau (1995).

This approach is only practical for the z dropout candidates and the main results are presented in Figure 7 (*rightmost panels*), where we overplot the best-fit templates on each SED and present the redshift probability distribution $P(z)$, marginalized over the entire parameter space (templates and reddening). As has been found by many workers (Lanzetta et al. 1996; Richard et al. 2006), the likelihood function reveals two peaks with different relative intensities, the relative height of the lower redshift peak acting as a valuable measure of foreground contamination, as it is linked to the color degeneracy between high-redshift star-forming galaxies and lower redshift early-type or reddened objects.



MOIRCS - K-band images of drop-outs

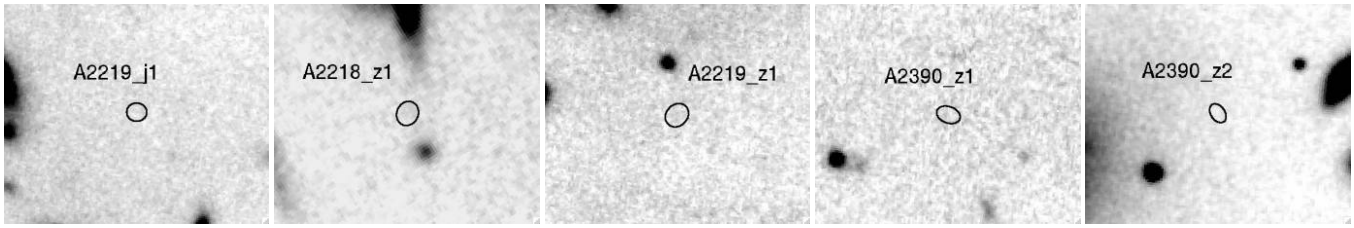


FIG. 7—*Continued*

Encouragingly, each candidate is more likely to be at high redshift and the probability of a foreground source is negligible for 4 out of our 10 sources. Integrated over all our candidates, we use $P(z)$ to compute the probability that each object lies beyond a redshift of 6, $\alpha_6 = P(z > 6.0)$, after normalizing $P(z)$ to unity over the redshift range $0 < z < 10$. We find α_6 values ranging from 0.46 to 0.91, with an average value of 0.65.

3.4.2. *Stacked Imaging*

As is evident from Table 3, the individual candidate SEDs are mostly restricted to detections in only one or two bands, with upper limits at other wavelengths. Although this precludes precise photometric redshift measures for each candidate, we can make further progress by combining data over several adjacent wave bands, and also by examining the average SED of the population to see if it is statistically consistent with that expected for a high-redshift source.

Several of our clusters have ACS data in multiple bands (Abell 2218, CL 1358, Cl 0024), which we combined after aligning the images with integer ACS pixel shifts and normalization each to a constant signal to noise ratio. The depth of this combined image, covering the wavelength range 4500–8000 Å, is typically 0.6–1.0 mag deeper than the individual bands. Yet in each of the

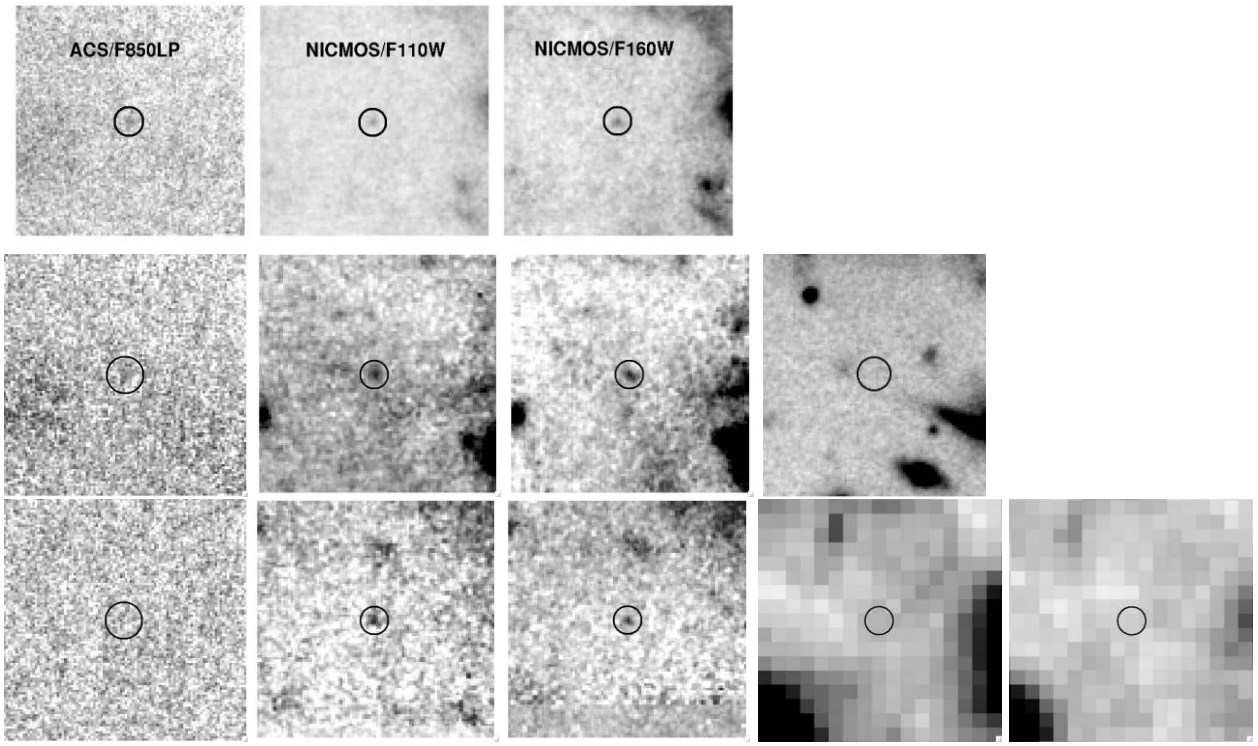


FIG. 8.—Stacked images of the *z*-band dropouts. *First row*: ACS and NICMOS images of all 10 sources. *Second row*: Stacked images for the four dropouts with *K*-band imaging (*rightmost image*). *Third row*: Stacked images for those six sources with unconfused IRAC data.

five *z* dropouts, no optical detection is seen, increasing our confidence level in the corresponding dropouts.

We likewise generated stacked *z*, *J*, and *H* images for all 10 *z* dropouts, as well as a stacked *K* image for the 4 sources observed with MOIRCS. In this case, We selected a $10'' \times 10''$ region around each candidate, and averaged the data rejecting 20% of the outlier pixels. For 6 objects free from contamination from nearby objects, the same stacking procedure was performed in the 3.6 and 4.5 μm IRAC bands. The stacked images are presented in Figure 8. No significant flux was detected in either the stacked MOIRCS or IRAC images.

Using the SExtractor parameters adopted for processing the individual images, an average SED was constructed for three populations depending on the availability of multiband ACS, *K*-band, and IRAC imaging. The photometric properties of each are listed in Table 4 and the SEDs are shown in Figure 9. These various stacked data sets offers a new opportunity to address the question of foreground contamination.

Using HyperZ, we fit these three SEDs exactly as described in § 3.4.1 Bolzonella et al.(2000). The optimal photometric redshift is consistent in each case (Fig. 10), with $z_{\rm ph} = 7.35 \pm 0.07$, and a reduced χ^2 lower than 1. We also fitted the same SEDs, but now restricting the photometric redshift to the range $0 < z_{\rm ph} < 3$. The best fit at low redshift is found at $z_{\rm ph} = 1.75 \pm 0.03$, but with a much higher χ^2, between 5 and 10. To the extent that the low-redshift solutions are credible, they imply galaxies with typically ages of 500–700 Myr and extinction values of $A_V \sim 1.0$–1.4.

One question that arises is whether the near-infrared $J - H$ color for our *z* dropouts is consistent with expectations, and also with that of the other limited detections at high redshift (Bouwens et al. 2008). Using the stacked images, we find this color is typically very blue: $J_{110} - H_{160} \sim 0.0$. We define the rest-frame ultraviolet slope β as $f_\lambda \propto \lambda^{-\beta}$ between the rest-frame wavelengths 1500 and 2000 Å (Calzetti et al. 1994), and estimate the uncertainty in β for our sources either from the dispersion within the range of plausible redshifts (at $\pm 1\sigma$), or by using the adopted photometric error bars in *J* and *H*. In both cases, we find a consistent value: $\beta \sim 2.8^{+0.05}_{-0.2}$. The mean slope is somewhat bluer than the average value of 2.0 ± 0.5 for a sample of *i*-band dropouts measured at $z \sim 6$ by Bouwens et al. (2006) but within the range of $\beta \sim 3.0$ found by Bouwens et al. (2004) in their sample of *z* dropouts, and marginally consistent with $\beta \sim 2.5$ given for several objects by Bouwens & Illingworth (2006). In order to produce such a blue ultraviolet slope, the SED can only be fit by very young models (typically <100 Myr) with little or no reddening ($A_V < 0.1$).

3.4.3. *Galactic Stars*

Our next test for foreground contamination concerns the question of low-mass Galactic stars. Various authors (Stanway et al.

TABLE 4
STACKED PHOTOMETRY OF THE *z*-BAND DROPOUTS

Dropouts	$z_{850\rm LP}$	$J_{110\rm W}$	$H_{160\rm W}$	K	IRAC$_{3.6\ \mu\rm m}$	IRAC$_{4.5\ \mu\rm m}$
All 10	28.59 ± 0.21	25.72 ± 0.14	25.70 ± 0.14	...	...	...
All four with *K* band	28.99 ± 0.32	25.71 ± 0.14	25.81 ± 0.12	>26.2	...	...
All six with IRAC	29.10 ± 0.23	26.11 ± 0.16	26.29 ± 0.16	...	>25.0	>24.8

NOTE.—Upper limits correspond to 5 σ.

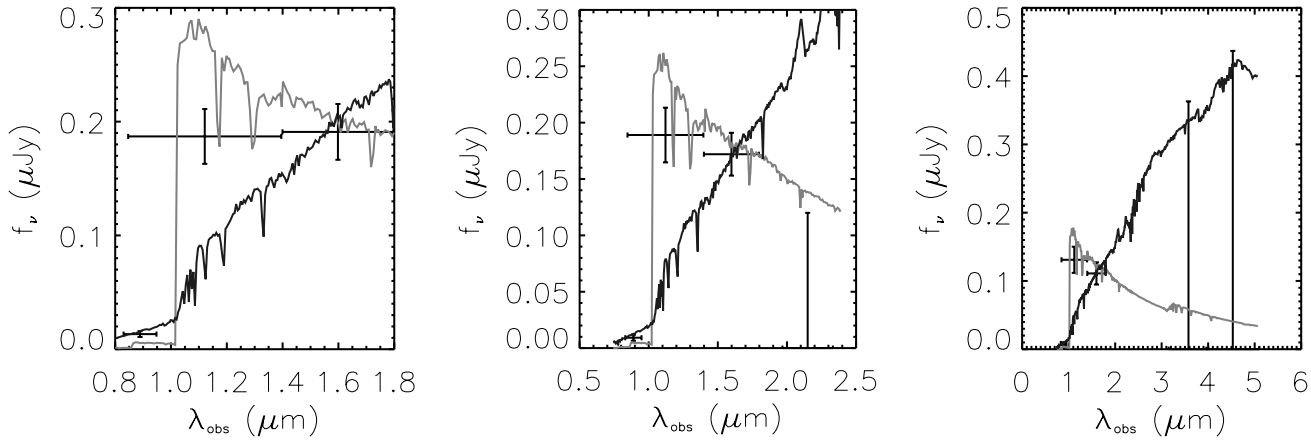


FIG. 9.—SED of the z-band dropouts derived from the stacked photometry (see Fig. 4). *Left to right*: All 10 dropouts with ACS/NICMOS data, dropouts with useful ACS/NICMOS/MOIRCS data, dropouts with useful ACS/NICMOS/IRAC data. In each case, the best template found with HyperZ over $0 < z < 10$ (*red curve*) or $0 < z < 3$ (*blue curve*) is shown. [*See the electronic edition of the Journal for a color version of this figure.*]

2004; Bouwens et al. 2004) have pointed out the difficulty of using optical–near-infrared colors to distinguish between cool stars and breaks arising from Lyman absorption at high redshift. L (Cruz et al. 2007) and T (Burgasser 2004) dwarfs exhibit metal and H_2O absorption features which produce features similar to the dropout signature. Indeed, L and T dwarfs from the SDSS and 2MASS surveys (Knapp et al. 2004) lie well within our color-color selection region (Fig. 3). Likewise, we find χ^2 values similar to the best template spectrum from HyperZ when fitting the SEDs of the candidates with a library of L0 and T dwarf templates spectra.

In such circumstances, *HST* data has occasionally been used to evaluate whether the sources have half-light radii R_h consistent with being extragalactic objects (Stanway et al. 2004). However, this is not a definitive criterion as we already know that many spectroscopically confirmed low-luminosity high-redshift sources are unresolved (Ellis et al. 2001). We measured R_h, defined as the radius enclosing half of the flux in the detection (NICMOS) images, and compared the values with that derived for bright not-saturated stars (0.2″). Figure 11 represents the location of our candidates in a R_h versus $J_{\rm AB}$ diagram, together with all other objects in our photometric catalog. Unfortunately, surface brightness dimming generates a cutoff at large values of R_h, which does not allow to distinguish resolved and resolved sources at the faintest magnitudes. This is the case for four of the z dropouts. At most we can say that two dropouts are unresolved, whereas four are resolved.

Noting the difficulty of separating stars from galaxies at these faint limits, a more practical approach is to examine the likely contamination statistically. Using simulations by Burgasser (2004), we computed the number of expected low-mass stars in the total area surveyed with NICMOS (7.7 arcmin^2), for all spectral types between L0 and T8, up to the extent of the thick disk (1 kpc). Assuming a slope $\alpha = 0.0$ for the mass function, consistent with recent observations by Metchev et al. (2008) we predict only 1.1 dwarf in our survey. A more pessimistic $\alpha = 0.5$ slope yields 1.5 stars. Thus, while there is undoubtedly some uncertainty surrounding these predictions, it does seem unlikely that cool Galactic stars represent a significant contaminant at such faint limits.

3.5. *Magnification and Multiple Images*

Our final test concerning the high-redshift nature of our candidate dropouts concerns their location within the image plane

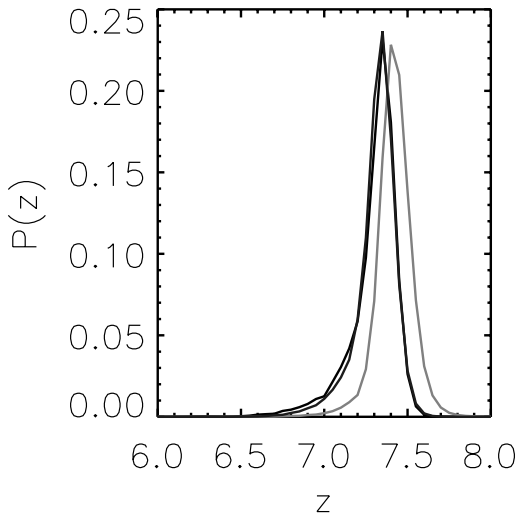


FIG. 10.—Redshift probability distribution estimated using HyperZ for the stacked SEDs of the z-band dropouts shown in Fig. 9. The black, red, and blue curves refer respectively to all 10 dropouts, those four with K-band imaging, and those six sources with useful IRAC upper limits. All three solutions are consistent with a mean population redshift of $z \sim 7.35$. [*See the electronic edition of the Journal for a color version of this figure.*]

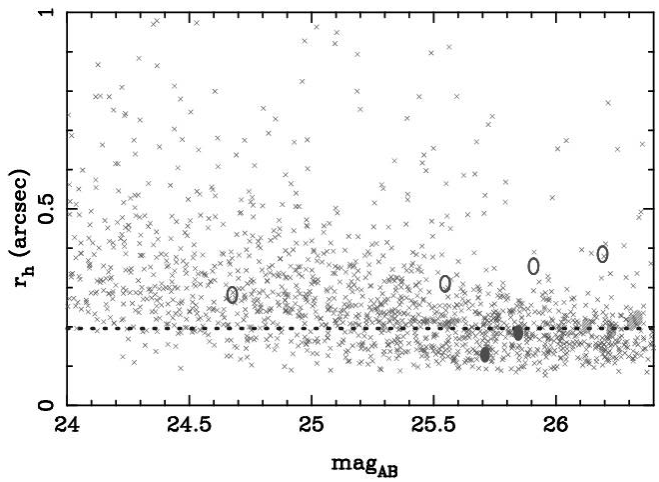


FIG. 11.—Distribution of half-light radii r_h measured by SExtractor in the photometric catalogs, as a function of the total J-band magnitude. The blue dashed curve corresponds to the measured r_h for bright nonsaturating stars. The 10 z-band dropouts are shown as red ellipses. Two objects are unresolved (*filled ellipses*), whereas four appear resolved (*open ellipses*). The rest cannot be reliably categorized. [*See the electronic edition of the Journal for a color version of this figure.*]

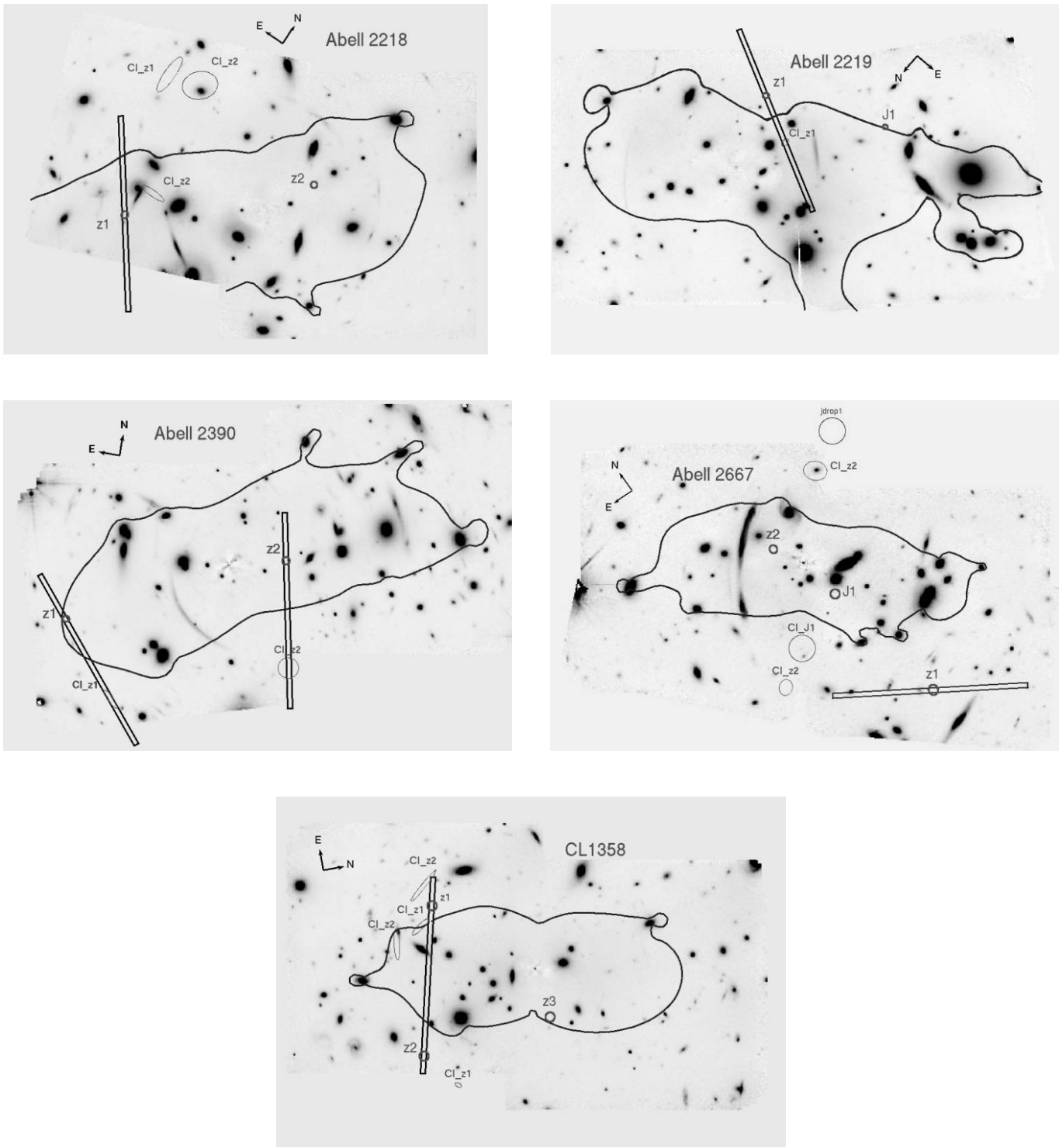


FIG. 12.—Location of the *z*-band and *J*-band dropouts with respect to the high-redshift ($z = 7.5$) critical line for each cluster field (*blue curve*). Ellipses with a "CI" label mark the position (and estimated error) of the brightest counterimages. The adopted NIRSPEC follow-up slit position angles are overplotted as black rectangles. [*See the electronic edition of the Journal for a color version of this figure.*]

of the lensing cluster. A key question is whether any might be expected to be multiply imaged as was the case, for example, for the object studied by Kneib et al. (2004). Most of the area covered by our NICMOS observations lies within the region of high magnification ($\mu > 2.5$ mag) where multiple images may occur.

Using the mass models for each cluster (updated using Lenstool from those referenced in Table 1), we estimate the magnification of each dropout based on its photometric redshift and location and examine the likelihood of any counter images as well as their predicted location and relative fluxes. The Bayesian optimization method incorporated in Lenstool also provides the uncertainties in these quantities.

In the majority of the cases (7 out of the 10 *z* dropouts), the model predicts a pair of images with similar fluxes (within 0.2–0.4 mag) straddling the critical line (Fig. 12). Single images are expected in two other cases (CL 1358-z3 and A2667-z1) and in

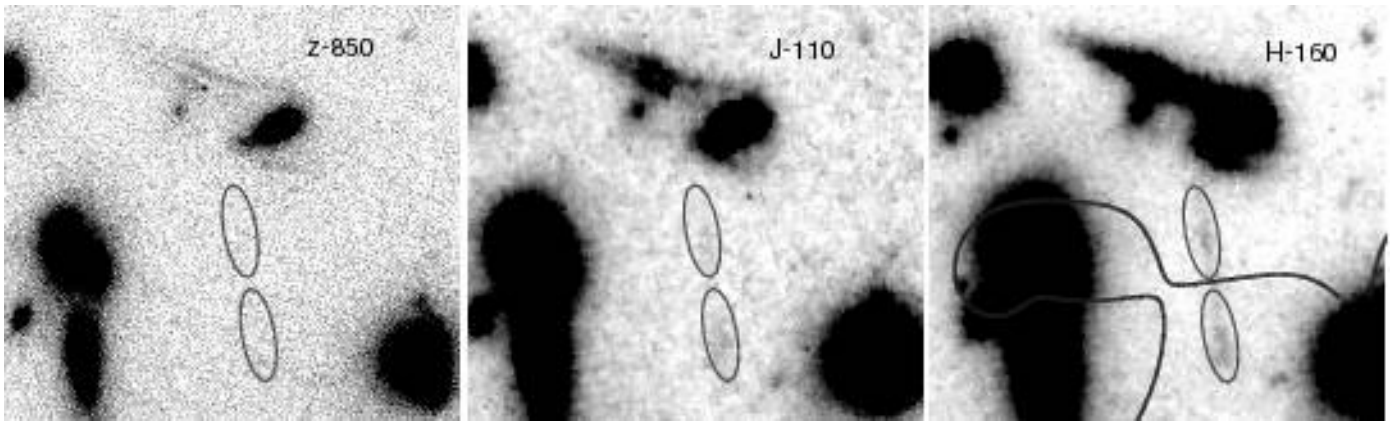


FIG. 13.—Pair of magnified optical dropouts identified in the NICMOS images. The theoretical location of the $z = 1.8$ critical line (*right panel*) confirms this source to be a low-z contaminant. [*See the electronic edition of the Journal for a color version of this figure*.]

the final case (CL 1358-z1), the objects sits on the predicted critical line, but is not expected to be viewed as a distinct pair at the finite angular resolution of NICMOS. These three cases are consistent with our observations. Considering the two *J* dropouts, A2219-j1 is predicted to be another example of close merging unresolved by *HST*, and A2667-j1 is predicted to have much fainter counterimages (by 0.8–1.5 mag), below our detection limits.

Our attention is thus focused on the seven cases where second images are expected. The typical positional errors are around 1″–3″. Unfortunately, most are located outside the area surveyed by NICMOS or in regions close to the edge of the detector or under the central BCG where the noise level is high. Only in two cases, A2218-z1 and A2667-z2, does the mass model predicts a detectable counterimage in a relatively clean region of the NICMOS detector. Unfortunately, no significant flux, within the range expected, was seen at either position in the *J*-band of *H*-band image.

Unfortunately, this test, valuable in principle, provides an inconclusive outcome in our exhaustive quest to confirm the high-redshift nature of our candidates. Although the high-redshift test is only applicable to two of our 10 *z*-band dropouts, we are unable to see either of the counter images. However, an identical search for multiple images was performed under the assumption that each candidate is at a lower redshift, $z \sim 2$. We would then expect counterimages at different locations, closer to the cluster center. However, we did not detect any of these predicted counter images either, for 4 clear cases. Only in a more marginal *J*-dropout candidate, which we dropped at an early stage, did this test succeed in demonstrating a $z \simeq 2$ solution (Fig. 13). Given the strong likelihood that most of the candidates lie either at $z \sim 2$ or $z \sim 7$–8, the test has a confused outcome. On the one hand, for those high-z candidates where we could expect to see a counterimage, none is seen. On the other hand for a larger sample of candidates, assumed to be at $z \simeq 2$, none is seen either. Accordingly, we deduce the test is not effective as a redshift discriminant.

3.6. *Summary*

We now summarize the possible success rate of our survey in generating high-redshift sources, concentrating on the *z*-band dropouts. Out of our 10 sources, we find that at most one source is spurious and ≃one is a Galactic star. Thus we conclude eight are likely to be extragalactic sources. Admittedly, about half of our candidates are unresolved, but we believe this is to be expected given the intrinsically faint limits we are probing with our lensing method. Discounting the inconclusive test based on counterimages (§ 3.5), and noting the 35% contamination from $z \simeq 2$ sources, we conclude that, statistically, we can expect 5 of our 10 sources to be $z > 6.8$ star-forming galaxies.

The foregoing analysis, while exhaustive, is however, statistical in nature. Our approach has been to treat all candidates as equally possible and to determine the level of foreground and spurious contamination as a fraction, without commenting on the nature of each individual source. Important information is contained in the similarity or otherwise of the candidate's morphology in the various NICMOS bands and the confidence with which we see no optical detection at the location of each candidate. For the *z*-band dropouts, a marginal detection is permitted in the red wings of the F850LP filter, but any hint of a signal at the location of the candidate in shorter wave bands would give cause for concern. It is then a matter of judgment whether to rely primarily on the photometric redshift solution (§ 3.4.1) or to override such information and reject a candidate after visual inspection.

Concentrating again on the *z*-band dropouts, two sources, A2219-z1 and A2390-z1, are resolved and satisfy the more rigorous color cut in Figure 3. One might therefore imagine these are particularly robust candidates. A2219-z1 has a detection in the WFPC2 F702W filter very close to the location of the NICMOS image. Although reasonably significant (27.1 or 4.5 σ), it does not influence the photometric solution (Fig. 6). Nonetheless, it does raise doubt about this candidate.

In the case of A2390-z1, there is no detectable optical signal down to F555W = 27.6 (2.0 σ) F814W = 27.0 (2.5 σ) and F850LP = 27.8 (2.0 σ), but the candidate's morphology differs somewhat between the NIC3 F110W and F160W filters. This may reflect the presence of two sources, one or both of which is at high redshift, or a genuine structural difference in the bands, for example as a result of line emission.

The only other source worth commenting on is A2667-z2 for which there is a marginal WFPC2 F606W detection at the position of the NIC3 source (28.25, or 2.0 σ). This is reflected in the fact that the photometric redshift solution is fairly ambiguous for this source (Fig. 6).

Finally, we note as in § 3.1 that the $J - H$ color of most of our *z* dropouts are significantly bluer than predicted for a normal SED at $z \simeq 7$. Shifting the sources to lower redshift would not significantly resolve this interesting observation.

Concerning the *J*-band dropouts, neither are particularly compelling. A2667-j1 has a similar morphological difference between

TABLE 5
Log of the Spectroscopic Observations Performed on Seven of the z-Band Dropouts

Run	Candidate(s)	Dither (arcsec)	Exposure Time (s)	Seeing (arcsec)	Notes
Jan 2007.....	CL 1358-z1 and z2	3	9 × 600	0.5	
2007 May.....	CL 1358-z1 and z2	3	15 × 600	0.5–0.6	
	A2219-z1	5	13 × 600	0.5–0.8	
			12 × 600	0.5	
	A2218-z1	5	18 × 600	0.9	Candidate + counterimage
			13 × 600	0.5	
	A2390-z1	5	4 × 600	0.8	
			12 × 600	0.5	
2007 Sep	A2390-z2	8	21 × 600	0.5	Candidate + counterimage
	A2667-z1	6	17 × 600	0.5	

Notes.—From left to right: epoch of observation, candidate name, spatial dithering between individual exposures, expoure time, seeing conditions. For two z dropouts, we managed to observe the predicted location of the counterimage at the same time as the candidate.

the F160W and faint F110W image and A2219-j1 has a marginal detection (27.0, 3.5 σ) in the ACS F850LP filter.

4. SPECTROSCOPIC FOLLOW-UP

Although deep imaging with *HST* has delivered candidates whose photometric redshifts lie beyond $z \simeq 7$, to date there has been not a single spectroscopic confirmation despite heroic efforts. A case in point is the $z \simeq 6.8$ lensed system in Abell 2218 (Kneib et al. 2004), which was the subject of 9.2 ks exposure with the LRIS optical spectrograph and a marathon 33 ks exposure with the NIRSPEC infrared spectrograph. A marginal continuum was seen but no emission lines were detected. This contrasts with the successful detection of Lyα in IOK-1 (Iye et al. 2006) in 31 ks. The latter source has an implied star formation rate of $\simeq 10\, M_\odot\ \mathrm{yr}^{-1}$ whereas the source in Abell 2218, when magnified, was expected to have an observed line flux equivalent to an unlensed system with a star formation rate of $(2.6 \times 25) \simeq 50\, M_\odot\ \mathrm{yr}^{-1}$. A tantalizing explanation for the nondetection of Lyα in the object in Abell 2218 is preferential damping by neutral hydrogen in lower luminosity sources.

The presence of the Lyα line provides a critical feature for confirming the nature of candidate high-redshift galaxies. However, the line is relatively easily attenuated and therefore may well be obscured in actively star-forming galaxies. Therefore, the absence of the line does not provide evidence that high-redshift candidates are false. Nevertheless, we conducted an ambitious spectroscopic campaign at Keck for some of our candidates. We naturally hoped that we might also secure the first spectroscopic verification of a $z > 7$ source.

4.1. *Observations*

We used the Near InfraRed SPECtrograph (NIRSPEC; McLean et al. 1998) on the 10 m Keck II Telescope to follow up the majority of our candidates in the window 0.964–1.120 μm, corresponding to the redshift range 6.9–8.2 for the Lyα line (1216 Å). Observations where conducted in several runs between 2007 January and September and we secured good data for seven z-band dropouts in total (see Table 5).

We used a 42″ long and 0.76″ wide slit, offering a resolving power of $R \sim 1500$ and used dithered exposures of 10 minutes each. We adjusted the dithering distance (in the range 3″–8″) in each case (third column of Table 5), to prevent overlap between a candidate and another bright source. Occasionally it was possible to observe two candidates simultaneously. For two candidates where we expect multiple images, the location of the expected counterimage (see § 3.5) was used to optimize the slit positions (Fig. 12).

The NIRSPEC spectra were reduced following the flat-fielding, sky subtraction, distortion corrections, and flux calibration procedures described by Stark et al. (2007b). These reduction techniques ensure an improved removal of the sky background by subtraction prior to resampling. We observed standard stars each night and used these to flux-calibrate the final spectra and determine the associated variance and hence the 5 σ limiting line flux. Each position was observed for about 3.5–4 hr in total, yield a limiting line flux of $\sim 3 \times 10^{-18}$ ergs cm^{-2} s^{-1} in regions of minimum OH contamination (Fig. 14), assuming a line width $\sigma_\alpha \sim 300$ km s^{-1}, as measured in the well-studied lensed system by Ellis et al. (2001). For a different line-width value, this sensitivity would vary as $(\sigma_\alpha)^{1/2}$.

We inspected each reduced two-dimensional spectrum for faint emission lines at the position of the candidate and, where relevant, that of the counterimage. No significant signal was detected for any of the candidates.

4.2. *Implications*

While the outcome of our spectroscopic campaign is certainly disappointing, the presence of strong OH lines in z band means

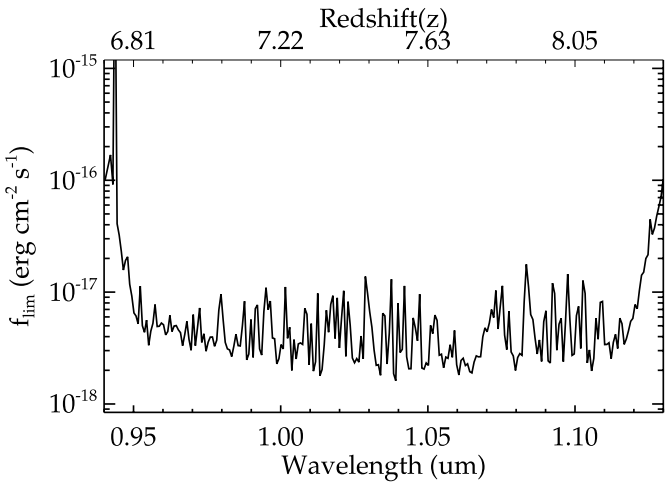


Fig. 14.—Determining the limiting flux from the NIRSPEC follow-up campaign. The plot shows the 5 σ limiting emission line flux vs. wavelength (and inferred Lyα redshift) for a typical 3.5 hr integration.

that our limiting Lyα flux ($\sim 3 \times 10^{-18}$ ergs cm^{-2} s^{-1}) applies only across 50% of the observed wavelength range. Thus we would only expect partial success even if all of our sources had intense emission lines. Nonetheless, it is informative to consider what the absence of any emission might mean given the star formation implied by our continuum detections. We can convert our flux limit into a rest-frame equivalent width using our *HST* photometry. For our candidates, we find a typical upper limit of $W_\lambda \sim 5$–20 Å.

Stanway et al. (2007) have recently studied a sample of faint Lyα emitters at $z \sim 6.0$ selected from their photometry to be i-band dropouts. They found a tail of high values for the W_λ distribution compared to similar studies undertaken at $z \sim 3.0$ (Shapley et al. 2003). They attribute this evolution to a tendency for stronger line emission in intrinsically faint sources. Fifty-eight percent of their sample has $W_\lambda > 25$ Å rest-frame. Including the lower and upper limit measurements of W_λ from this sample, which most likely contains lower redshift contaminants, the fraction is lowered to 34%.

Assuming no evolution in this distribution from $z \sim 6$ to $z \simeq 7$–8, we would expect in the optimistic case (58% value) about four Lyα emitters in our spectroscopic campaign, prior to considerations of the OH spectrum. The probability that all four objects lie in a region of the spectrum contaminated by OH emission is thus $(0.5)^4 \sim 0.06$, which is low. Even if only four of the seven candidates we examined were at high redshift (based on our statistical estimates given in § 3), we should expect to detect the emission for $\simeq$3 cases. Here there would be only a 12% probability of each one being occulted by OH emission. In the most pessimistic case from Stanway et al. 2007 (34%), we would expect to detect only one source, with 50% probability of OH contamination.

Thus, as in the case of the $z \simeq 6.8$ source in Abell 2218, the absence of emission in 7 candidates is somewhat surprising. Assuming a significant fraction are at high redshift as discussed in § 3, this may be an important indication of the evolution in the intergalactic medium above $z \simeq 6$. Regardless of the cause, it adds to the challenge of making progress in verifying high-redshift candidates.

Reconciling the above with the abundance of intrinsically faint Lyα emitters claimed by Stark et al. (2007b) is admittedly difficult. Should the bulk of the dropout population at $z > 7$ continue to reveal no emission, this would suggest a moderate neutral fraction that would challenge the transparency of the IGM at $z \simeq 8$–10 implied by the presence of feeble Lyα emitters. The enigma simply reinforces the importance of continuing to attempt the detection and verification of line emission in very faint sources.

5. DISCUSSION

In the foregoing we have described a concerted effort to quantify the abundance of low-luminosity star-forming galaxies conducted in parallel to a similar spectroscopic campaign which has examined the abundance of $z \simeq 8$–10 Lyα emitters (Stark et al. 2007b). That study claimed that if even a small fraction of the candidates is truly at high redshift, a significant contribution to reionization is provided by low-luminosity galaxies. In a similar manner, recognizing the limitations of our small samples, we now examine the luminosity function at $z \simeq 7.5$ and the possible contribution that our lensed dropout sources may make to cosmic reionization.

5.1. *Number Densities and the Source Luminosity Function*

The intrinsic area of sky (i.e., that in the source plane) covered by the NICMOS images is strongly dependent on the geometry of the critical lines (or caustics), which varies from cluster to cluster. Furthermore, multiple images occur in the central regions,

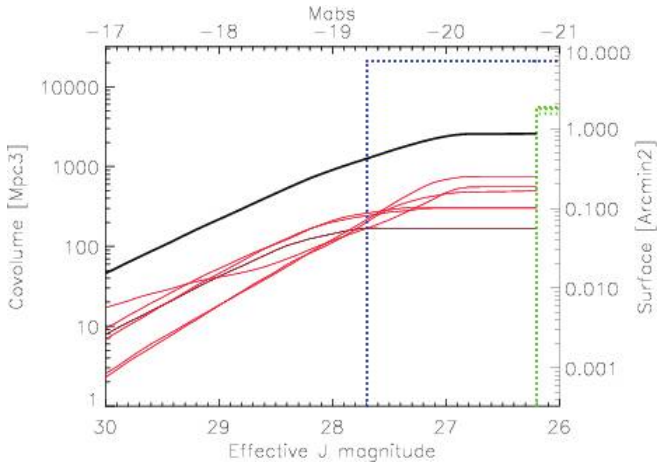


FIG. 15.—Survey characteristics: the intrinsic (unlensed) surface area sampled in the source plane down to a given limiting magnitude for each cluster (*thin red lines*) and for all six clusters (*thick black line*). The upper scales give the corresponding absolute magnitudes assuming $z \sim 7.5$. The green dashed lines illustrate the areas sampled in the absence of lensing. The blue dotted line shows the equivalent survey parameters for the UDF (Bouwens et al. 2004; Bouwens & Illingworth 2006).

duplicating the corresponding source plane area. The result of both effects is a smaller survey, reduced by a typical factor of $\simeq$10 in the source plane, with an increased depth whose value varies across the field of view.

In order to derive the source density of our z-band dropouts, and to compare our results with those conducted in blank fields, we used the lensing models for each cluster to compute the sky area effectively observed in the source plane, down to a given intrinsic AB magnitude. We assumed our survey covers the image plane down to the measured 5 σ depth $J_{110} \sim 26.2$ in the central NICMOS region, and scaled according to the relative exposure time near the edges. We also removed $\sim$10% of the NICMOS area affected by bright galaxies. We supposed an average redshift of $z = 7.5$ to compute the magnification factors.

Errors in the magnification factors estimated from the lensing models will affect the source plane areas and unlensed magnitudes in an opposite way. For an individual cluster, the typical uncertainty is about 10% in the resulting area. This error is even smaller for the entire sample of six clusters.

Our total surveyed area is a factor $\sim$5 smaller than the UDF in the same magnitude range (AB < 27.7, Fig. 15). However, the increased depth enables to reach AB ~ 28–30 in this area. A very similar result is found in the case of J-band dropouts, assuming $z = 9$.

We used the estimated color selection contamination factor, $f_{\rm cont}$, and the selection completeness factor, $f_{\rm comp}$ (§ 3.1 and Table 3), to correct each z dropout individually to derive intrinsic number densities. Because of the strong variations in the magnification factor across the NICMOS field of view, we corrected for observed completeness by computing the completeness factors in intrinsic (unlensed) $\langle J + H \rangle$ magnitude. This combines both the observed completeness factor, $S_{\rm comp}$, given in Table 2 and the surface reduction in the source plane. Error bars in the number densities were computed using Poisson noise estimates. We present the cumulative UV luminosity function of the z-band dropouts in the magnitude range $27.0 <$ AB < 30.0 in Figure 16.

As discussed, it is likely that 5 out of our 10 sources are at high redshift. Accordingly, in Figure 16, we randomly selected 100×5 sources from our sample to take into account object-to-object variations in the magnification factor and used this to estimate

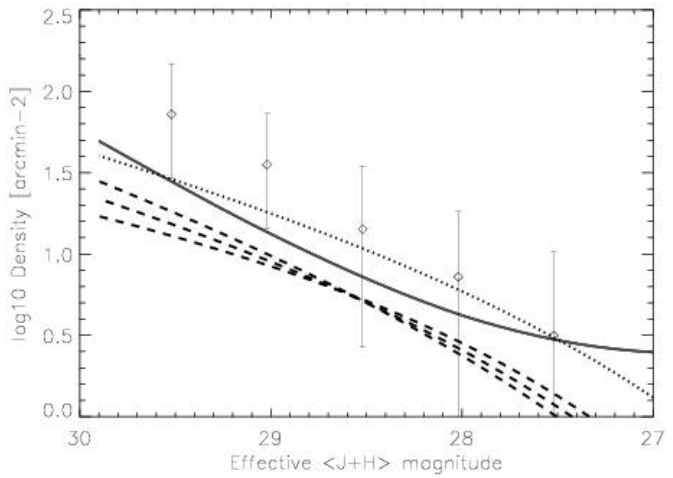


FIG. 16.—Cumulative number density of star-forming galaxies at $z \sim 7$–8 as a function of the effective (unlensed) magnitude. Data points and error bars correspond to the range of densities resulting when randomly selecting five candidates from our sample and adopting Poisson errors (*open diamonds, offset for clarity*). In the most pessimistic case, where no sources are at high redshift, the implied upper limit is shown by the thick red curve. We overplot the best-fit luminosity functions found by Bouwens et al. (2006; *light dotted line*) and Bouwens et al. (2008; *heavy dashed lines*) in the UDF (parameters in Table 6). [*See the electronic edition of the Journal for a color version of this figure.*]

more realistic error bars. Errors in the individual magnification factors vary between 0.05 and 0.2 mag. When compared with the luminosity bins we used (0.5 mag), this effect has very low significance on the results, and the errors are mainly recovered when randomly choosing five candidates from the sample.

For comparison, we overplot in Figure 16 the luminosity function found by Bouwens et al. (2006) in the UDF, including redshift evolution between $z = 6.0$ and $z = 7.5$ assuming the observed size scaling as $(1 + z)^{-1}$ for fixed luminosity (Ferguson et al. 2004). Likewise, we overplot the best Schechter function fits recently claimed by Bouwens et al. (2008) from an analysis of their sample of z-band dropouts. Not surprisingly, there is no overlap between these blank field measures and our, much deeper, lensed survey. All that can be said is that our results, which probe more than $\simeq$2.5 mag fainter are marginally consistent down to AB $\sim$ 28.5, and higher by $\sim$0.3–0.6 dex at fainter luminosities. Likewise, we overplot the best Schechter function fits recently claimed by Bouwens et al. 2008 from an analysis of their sample of z-band drop-outs (Table 6).

5.2. *Contribution to Cosmic Reionization*

We finally investigate whether the likely abundance of low-luminosity sources found in our survey could make a significant contribution to cosmic reionization. The approach we use is somewhat similar to the one described by Stark et al. (2007b), who estimated the comoving number density of sources necessary to keep the intergalactic medium (IGM) reionized under reasonable assumptions, and compare those to the abundances derived from candidate Lyα emitters at high redshift.

In our case, we can estimate the star formation rate density, measured in individual objects from their UV rest-frame luminosity, after applying the same completeness corrections described in § 5.1. We converted the intrinsic (unlensed) $J + H$ magnitudes of our candidates into a UV luminosity, L_{1500}, and infer the related star formation rate (SFR) by adopting the Kennicutt (1998) calibration. All z-band dropouts span the range SFR $\sim$ 0.1–1.0 $M_\odot$ yr^{-1}, thus the overall star formation rate observed yields the contribution of low star formation rate sources to the entire star formation rate density $\rho_{\rm SFR}$.

TABLE 6
BEST-FIT SCHECHTER PARAMETERS OF THE HIGH-REDSHIFT UV LUMINOSITY FUNCTION

Reference	z	Φ^* (Mpc^{-3})	M^* (mag)	α
Bouwens et al. (2006)	6	1.4e-3	−20.25	−1.73
Bouwens et al. (2008)	7.4	1.1e-3	−19.80	−1.74
		1.7.8e-3	−19.60	−1.4
		8e-4	−19.90	−2.0

NOTES.—From earlier results found in the UDF, and overplotted in Fig. 16. For each work we give the normalization Φ^*, the absolute magnitude M^* at the exponential cutoff and the faint-end slope α. Bouwens et al. (2008) explore three possible evolutions of the L^* and Φ^* parameters for different fixed slopes α.

We used the Madau et al. (1999) formalism to estimate the amount of star formation necessary to keep the IGM reionized at a given redshift. One important factor in this calculation, that would modify the efficiency of star-forming sources to reionize the IGM, is the H II clumping factor C, defined as $C = \langle n^2_{\rm H\,II}\rangle/\langle n_{\rm H\,II}\rangle^2$ with $N_{\rm H\,II}$ being the density of ionized hydrogen. This factor measures the inhomogeneity of ionized hydrogen in the IGM which will likely increase between $z = 10$ and $z = 6$ due to the growth of structure. Assuming an IMF with a Salpeter slope with stellar masses ranging from 1 to 100 $M_\odot$, and a solar metallicity $Z = 0.02$, the photon budget from star-forming sources necessary to reionize the IGM can be written as

$$\dot{\rho}_{\rm SFR} \simeq \left(0.031\, M_\odot\ {\rm yr}^{-1}\ {\rm Mpc}^{-3}\right)\left(\frac{f_{\rm esc,rel}}{0.5}\right)^{-1}\left(\frac{C}{10}\right)\left(\frac{1+z}{8.5}\right)^3, \quad (1)$$

where $f_{\rm esc,rel}$ is the escape fraction of ionizing photons. We assumed an escape fraction $f_{\rm esc,rel} = 0.5$ in our calculations. However, values as low as $f_{\rm esc,rel} = 0.02$ have been measured in $z \sim 3$ galaxies by Shapley et al. (2006). Lower escape fractions would increase the amount of star formation necessary to reionize the universe, so adopting $f_{\rm esc,rel} = 0.5$ gives us a lower limit on $\rho_{\rm SFR}$. On the other hand, top-heavy IMFs and differences in metallicities would make galaxies produce more ionizing photons per star formation rate, but this effect is less significant than variations in $f_{\rm esc,rel}$ and C. Bolton & Haehnelt (2007) have critically reviewed possible values for the clumping factor C. Many authors (Bunker et al. 2004; Bouwens et al. 2006) have assumed $C = 30$, but much lower values are predicted from radiative transfer simulations (Iliev et al. 2006 find $C < 2$ at $z > 11$).

Figure 17 illustrates the star formation rate densities obtained by integrating down to a given SFR for the two extreme luminosity functions derived by Bouwens et al. (2008) when fitting their number densities at higher luminosities (equivalent to SFR $>$ 1.0 $M_\odot$ yr^{-1}). We overplot on this figure the contributions derived from equation (1) with clumping factor varying between $C = 2$ and $C = 30$.

As suggested before, luminous galaxies do not produce enough star formation to reionize the IGM at these redshifts, even when a low $C = 2$ clumping factor is assumed. Our sample enables us to test whether lower luminosity galaxies help to solve this discrepancy. Combining the source density from our NICMOS survey with those claimed at SFR $> 1.0\, M_\odot$ estimated by integrating

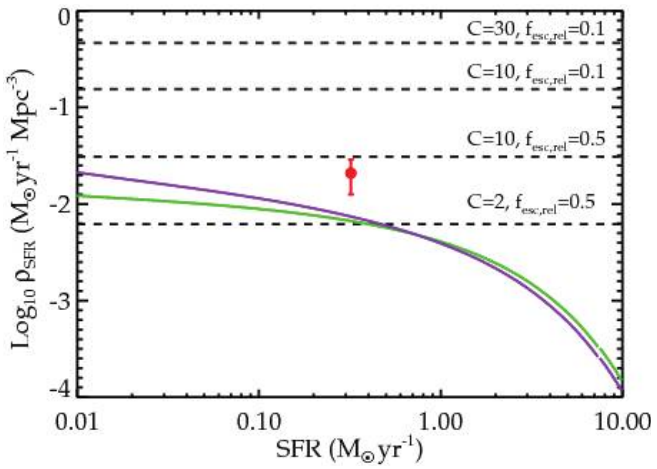


FIG. 17.—Cumulative comoving density of star formation rate at $z \simeq 7.5$ derived for the two extreme (two last entries of Table 6) luminosity functions from Bouwens et al. (2008) with faint end slopes of $\alpha = 2.0$ (*purple*) or $\alpha = 1.4$ (*green*). The constraints from the present survey are shown as a red circle, for the average and range of densities resulting when randomly selecting five candidates from our sample (as in Fig. 16). The density necessary to keep the IGM reionized at $z = 7.5$, calculated from eq. (1) for a range of clumping factors C and escape fraction $f_{\rm escp,rel}$, is shown as the dashed lines.

the luminosity function from Bouwens et al. (2008) we can compare the photon budget down to SFR $\sim 0.1\ M_\odot\ {\rm yr}^{-1}$ with the amount of star formation rate necessary to reionize the IGM. The results, shown in the case of five objects randomly chosen from our sample of 10 z-dropouts, suggest a contribution compatible with that necessary for reionization for clumping factors in the range $2 < C < 10$.

6. CONCLUSIONS

The overall goal of this project has been to constrain the abundance of low-luminosity star-forming galaxies at $z \sim 7$–10, selected as z- and J-band dropouts in the fields of six lensing clusters observed with ACS and NICMOS on board the *Hubble Space Telescope*, and the IRAC camera on board the *Spitzer Space Telescope*. We summarize our results as follows:

1. We have identified 12 high-redshift candidates (10 z-band dropouts and two J-band dropouts) according to carefully determined photometric selection criteria. These are located in five of the six clusters and span the observed magnitude range $J_{110} \simeq 25$–26. Each is typically magnified by 1.5 to 4 mag.

2. Based on a comprehensive set of tests, we estimate the fraction of sources that might represent spurious detections and the extent to which low-mass stars and low-redshift interlopers may contaminate our sample. Collectively, these tests suggest that around five of our 10 z dropouts are possible high-redshift $z > 7$ objects.

3. By stacking the available ACS, NICMOS, ground-based K-band, and IRAC images, we investigated further the averaged properties of our lensed z dropouts. We find a UV spectral slope $\beta \sim 2.8^{+0.05}_{-0.2}$ similar to that of higher luminosity candidates from the UDF. Such a slope suggests a very young stellar population with little reddening and strengthens our case that the bulk of our candidates are high-redshift sources.

4. We searched for possible counterimages for our candidates based on the most recent lensing models for each cluster. Unfortunately, our results are inconclusive. Many of the counterimages either lie outside our NICMOS field or are close to foreground sources. We fail to detect a counterimage in two apparently clean cases, but a further two sources may be potentially merging on the critical line.

5. We undertook follow-up spectroscopy with NIRSPEC for seven of our 10 z-band dropout candidates in the hope of seeing confirmatory Lyα emission. No emission was found in any candidate (or its counterimage location) to a flux limit corresponding to 3×10^{-18} ergs s^{-1} cm^{-2} in the clean part of the OH spectrum. One explanation is possible evolution in the Lyα rest-frame equivalent width distribution, compared to previous results by Stanway et al. (2007) at $z \sim 6$, such as might be expected if the neutral fraction rises with redshift. Such a deduction would be difficult to reconcile with the presence of intrinsically faint lensed Lyα emitters at $z \simeq 10$ Stark et al. 2007b.

6. Our inferred luminosity function at $z \sim 7.5$, after correcting for contamination and incompleteness, is marginally consistent with an extrapolation of available constraints at brighter luminosities, with a slightly higher normalization by 0.3–0.6 dex. If even a modest fraction of our sources are at high redshift, our results strengthen the suggestion that sources with star formation rates $\sim$0.1–1.0 $M_\odot$ yr^{-1} contribute significantly to cosmic reionization.

As we approach the era of *James Webb Space Telescope* and the Extremely Large Telescopes, the outcome of our project in lensed fields has been to provide new constraints on the faint part of the luminosity function at $z > 7$, which confirmed the trends seen at higher luminosities. Despite being restricted to a small field after demagnification in the source plane, we expect that these results will be readily confirmed by the upcoming Wide Field Camera (WFC3) on *HST*, combining extremely sensitive infrared channels with a field of view much larger than NICMOS.

We thank the anonymous referee for a report that has improved the content of this paper, and Rychard Bouwens for his very helpful comments on an earlier version of the manuscript. We also acknowledge discussions with Rodger Thompson, Elizabeth Stanway, Roser Pelló, Daniel Schaerer, Kelle Cruz, and Adam Kraus. Andrew Blain and Naveen Reddy kindly provided the Keck NIRC observations of Abell 2218. We are thankful to Ichi Tanaka for his support in the reduction of MOIRCS imaging data. We acknowledge funding from NASA grant HST-GO-10504.01-A and *Spitzer* program GO-20439. The authors recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. This program is based on observations made with the NASA/ESA *Hubble Space Telescope*, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555, the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan, the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA, and the Canada-France-Hawaii Telescope (CFHT), which is operated by the National Research Council of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii.

REFERENCES

Bardeau, S., Soucail, G., Kneib, J.-P., Czoske, O., Ebeling, H., Hudelot, P., Smail, I., & Smith, G. P. 2007, A&A, 470, 449
Beckwith, S. V. W., et al. 2006, AJ, 132, 1729
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bolton, J. S., & Haehnelt, M. G. 2007, MNRAS, 382, 325
Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476
Bouwens, R. J., & Illingworth, G. D. 2006, Nature, 443, 189
Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., & Franx, M. 2006, ApJ, 653, 53
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2008, ApJ, in press (arXiv: 0803.0548)
Bouwens, R. J., Illingworth, G. D., Thompson, R. I., & Franx, M. 2005, ApJ, 624, L5
Bouwens, R. J., et al. 2004, ApJ, 616, L79
Bradley, L. D., et al. 2008, ApJ, 648, 647
Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, ApJ, 438, 49
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bunker, A. J., Stanway, E. R., Ellis, R. S., & McMahon, R. G. 2004, MNRAS, 355, 374
Burgasser, A. J. 2004, ApJS, 155, 191
Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, ApJ, 533, 682
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Casertano, S. 2000, AJ, 120, 2747
Choudhury, T. R., & Ferrara, A. 2007, MNRAS, 380, L6
Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393
Covone, G., Kneib, J.-P., Soucail, G., Richard, J., Jullo, E., & Ebeling, H. 2006, A&A, 456, 409
Cruz, K. L., et al. 2007, AJ, 133, 439
Daou, D., & Skinner, C. 1997, in 1997 *HST* Calibration Workshop with a New Generation of Instruments, ed. S. Casertano et al. (Baltimore: STScI), 263
Dickinson, M., Giavalisco, M., & The Goods Team. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini (Berlin: Springer), 324
Egami, E., et al. 2005, ApJ, 618, L5
Elíasdóttir, Á. et al. 2007, preprint (arXiv: 0710.5636)
Ellis, R., Santos, M. R., Kneib, J., & Kuijken, K. 2001, ApJ, 560, L119
Eyles, L. P., Bunker, A. J., Ellis, R. S., Lacy, M., Stanway, E. R., Stark, D. P., & Chiu, K. 2007, MNRAS, 374, 910
Fazio, G. G., et al. 2004, ApJS, 154, 10
Ferguson, H. C., et al. 2004, ApJ, 600, L107
Franx, M., Illingworth, G. D., Kelson, D. D., van Dokkum, P. G., & Tran, K.-V. 1997, ApJ, 486, L75
Henry, A. L., Malkan, M. A., Colbert, J. W., Siana, B., Teplitz, H. I., McCarthy, P., & Yan, L. 2007, ApJ, 656, L1
Ichikawa, T., et al. 2006, Proc. SPIE, 6269, 626916
Iliev, I. T., Mellema, G., Pen, U.-L., Merz, H., Shapiro, P. R., & Alvarez, M. A. 2006, MNRAS, 369, 1625
Iye, M., et al. 2006, Nature, 443, 186
Jullo, E., Kneib, J.-P., Limousin, M., Elíasdóttir, Á., Marshall, P. J., & Verdugo, T. 2007, New J. Phys., 9, 447
Kashikawa, N., et al. 2006, ApJ, 648, 7
Kennicutt, R. C. 1998, ApJ, 498, 541
Kinney, A. L., Calzetti, D., Bohlin, R. C., McQuade, K., Storchi-Bergmann, T., & Schmitt, H. R. 1996, ApJ, 467, 38
Knapp, G. R., et al. 2004, AJ, 127, 3553
Kneib, J.-P., Ellis, R. S., Santos, M. R., & Richard, J. 2004, ApJ, 607, 697
Kneib, J.-P., et al. 2003, ApJ, 598, 804
Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2002, in 2002 *HST* Calibration Workshop: Hubble after the Installation of the ACS and the NICMOS Cooling System, ed. S. Arribas, A. Koekemoer & B. Whitmore (Baltimore: STScI)., 337
Komatsu, E., et al. 2008, preprint (arXiv: 0803.0547)
Lanzetta, K. M., Yahil, A., & Fernández-Soto, A. 1996, Nature, 381, 759
Limousin, M., et al. 2007, ApJ, 668, 643
Loeb, A., & Barkana, R. 2001, ARA&A, 39, 19
Madau, P. 1995, ApJ, 441, 18
Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648
McLean, I. S., et al. 1998, Proc. SPIE, 3354, 566
Metchev, S. A., Kirkpatrick, J. D., Berriman, G. B., & Looper, D. 2008, ApJ, 676, 1281
Oke, J. B. 1974, ApJS, 27, 21
Richard, J., Pelló, R., Schaerer, D., Le Borgne, J.-F., & Kneib, J.-P. 2006, A&A, 456, 861
Richard, J., et al. 2007, ApJ, 662, 781
Ryan-Weber, E. V., Pettini, M., & Madau, P. 2006, MNRAS, 371, L78
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K. 2006, ApJ, 651, 688
Shimasaku, K., et al. 2006, PASJ, 58, 313
Smith, G. P., Kneib, J.-P., Smail, I., Mazzotta, P., Ebeling, H., & Czoske, O. 2005, MNRAS, 359, 417
Songaila, A. 2004, AJ, 127, 2598
Spergel, D. N., et al. 2007, ApJS, 170, 377
Stanway, E. R., Bunker, A. J., McMahon, R. G., Ellis, R. S., Treu, T., & McCarthy, P. J. 2004, ApJ, 607, 704
Stanway, E. R., et al. 2007, MNRAS, 376, 727
Stark, D. P., Bunker, A. J., Ellis, R. S., Eyles, L. P., & Lacy, M. 2007a, ApJ, 659, 84
Stark, D. P., Ellis, R. S., Richard, J., Kneib, J.-P., Smith, G. P., & Santos, M. R. 2007b, ApJ, 663, 10
Stark, D. P., Loeb, A., & Ellis, R. S. 2007c, ApJ, 668, 627
Swinbank, A. M., Bower, R. G., Smith, G. P., Smail, I., Kneib, J.-P., Ellis, R. S., Stark, D. P., & Bunker, A. J. 2006, MNRAS, 368, 1631
van Dokkum, P. G. 2001, PASP, 113, 1420
Werner, M. W., et al. 2004, ApJS, 154, 1